Attorney's Docket No. 5800-49(35800/184745) PATENT

In The United States Patent And Trademark Office

In re:	Meyers	Confirmation No.:	7067
Appl. No.:	09/464,039	Group Art Unit:	1636
Filed:	December 15, 1999	Examiner:	Sumesh Kaushal
For:	21612 NOVEL HUMAN DEHYDROGENASE (Title as Amended 11 19/01)		

Box AF
Commissioner for Patents
Washington, DC 20231

APPEAL BRIEF TRANSMITTAL
(PATENT APPLICATION – 37 C.F.R. § 1.192)

1. Transmitted herewith, in **triplicate**, is the APPEAL BRIEF in this application, with respect to the Notice of Appeal filed on June 25, 2002.

2. This application is filed on behalf of
   - ☒ other than a small entity
   - ☐ a small entity

3. Pursuant to 37 C.F.R. § 1.17(c), the fee for filing the Appeal Brief is:
   - ☐ small entity $160.00
   - ☒ other than small entity $320.00

   Appeal Brief fee due $320

   ☒ Any additional fee or refund may be charged to Deposit Account 16-0605.

Respectfully submitted,

Kathryn L. Coulter

Kathryn L. Coulter
Registration No. 45,889

CUSTOMER NO. 00826
ALSTON & BIRD LLP
Bank of America Plaza
101 South Tryon Street, Suite 4000
Charlotte, NC 28280-4000
Tel Raleigh Office (919) 862-2200
Fax Raleigh Office (919) 862-2260

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to Box AF, Commissioner for Patents, Washington, DC 20231 on **September 24, 2002**.

Nora C. Martinez

Nora C. Martinez

Copy 1

Attorney's Docket No. 5800-49 (35800/184745) PATENT

#18

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re:	Meyers	Confirmation No.:	7067
Appl. No.:	09/464,039	Group Art Unit:	1636
Filed:	December 15, 1999	Examiner:	Sumesh Kaushal
For:	21612, NOVEL HUMAN DEHYDROGENASE [TITLE AS AMENDED 11/19/01]		

September 24, 2002

Box AF
Commissioner for Patents
Washington, DC 20231

APPEAL BRIEF

Sir:

This Appeal Brief is filed pursuant to the "Notice of Appeal to the Board of Patent Appeals and Interferences" mailed June 21, 2002 and received by the Office on June 25, 2002.

Real Party in Interest.

The real party in interest in this appeal is Millennium Pharmaceuticals, Inc., the assignee of the above-referenced patent application.

Related Appeals and Interferences.

There are no related appeals and/or interferences involving this application or its subject matter.

Status of Claims.

Claims 63-67, 77-79, and 87-104 are the subject of this appeal. The claims appear in Appendix A. Claims 1-62, 68-76, and 80-86 have been cancelled.

Status of Amendments.

All of Applicant's amendments have been entered.

Summary of the Invention.

The pending claims of the present invention are drawn to isolated nucleic acid molecules having a nucleotide sequence encoding the 21612 short chain dehydrogenase, nucleotide sequences encoding fragments of these polypeptides, and nucleotide sequences encoding functional variants of the 21612 short chain dehydrogenase, as well as host cells containing these nucleic acid molecules, methods and kits for detecting the 21612 nucleotide sequences, and methods of producing the encoded 21612 polypeptides, fragments, and variants. The 21612 dehydrogenase gives a high score when compared to a consensus sequence for alcohol dehydrogenases.

Issues.

Issue 1--Whether the invention of claims 63-67, 77-79, and 87-104 has utility under 35 U.S.C. §101.

Issue 2--Whether the invention of claims 63-67, 77-79, and 87-104 is enabled under 35 U.S.C. § 112, first paragraph.

Issue 3--Whether a sufficient written description of the invention of claims 88-92 is provided under 35 U.S.C. § 112, first paragraph.

Issue 4--Whether claim 79 is indefinite under 35 U.S.C. § 112, second paragraph.

Grouping of Claims.

The claims do not stand or fall together. While all the pending claims have been rejected under 35 U.S.C. § 101 and under 35 U.S.C. § 112, first paragraph, on the grounds that the claimed invention lacks utility and thus are not enabled, claims 88-92 have also been rejected under 35 U.S.C. § 112, first paragraph, on the grounds that the specification does not provide sufficient written description for the claimed subject matter. Furthermore, claim 79 has been rejected under 35 U.S.C. § 112, second paragraph, on the grounds that it is indefinite. Accordingly, the issues surrounding the claims are different, and the claims do not stand or fall together.

Argument.

Issue 1--Whether the invention of claims 63-67, 77-79, and 87-104 has utility under 35 U.S.C. §101.

The Examiner has rejected claims 63-67, 77-79, and 87-104 under 35 U.S.C. §101 as lacking utility. The Examiner argues that the claimed invention lacks patentable utility because "[t]he specification fails to show a single working example that establishes that the SEQ ID NO:8 is a member of Alcohol dehydrogenase (ADH) family, such as by any substantial sequence homology and/or functional assay of the protein." March 25, 2002 Office Action, page 3. In fact, Applicant has determined that the 21612 protein functions as an alcohol dehydrogenase using methods that are reliable and are accepted by those of skill in the art.
Accordingly, the claimed invention has a utility that is specific, substantial, and credible and therefore meets the requirements for patentability.

I. The 21612 polypeptide is an ADH short chain dehydrogenase

A. Pfam analysis demonstrates that the 21612 polypeptide is an ADH short chain dehydrogenase.

The function of the novel 21612 polypeptide was determined by comparing the 21612

amino acid sequence set forth in SEQ ID NO:7 to the Pfam database of protein families. In this analysis, the 21612 polypeptide gave a high score when compared to the Pfam consensus sequence for ADH short chain dehydrogenases (Pfam Accession No. PF00106). A copy of the alignment of the 21612 polypeptide with the Pfam short chain dehydrogenase consensus sequence is provided as Appendix B.

The Pfam database provides a curated collection of well-characterized protein family domains with high quality alignments. Functional domains of novel proteins may be identified by comparison with the Pfam protein family models. It is well known in the art that the presence of a consensus domain characteristic of a family of proteins having a know function may be used to determine the function of a novel polypeptide. The sequences included in the Pfam seed alignment used to create the short chain dehydrogenase consensus include proteins that have been well-characterized biochemically; for example an alcohol dehydrogenase from *Drosophila* (NCBI Accession No. P21898), several human estradiol 17 β-dehydrogenases (NCBI Accession Nos. P37059, P51659, and P14061), a human corticosteroid 11-β-dehydrogenase (NCBI Accession No. P80365), and a human 15-hydroxyprotaglandin dehydrogenase (NCBI Accession No. P15428).

The Examiner has stated that the Applicant's arguments regarding the reliability of the Pfam analysis are found unpersuasive "because PFAM analysis revealed that 21612 matches with a top-scoring domain for ADH-short but with low sequence similarity." March 25, 2002 Office Action, page 2, and August 14, 2002 Office Action, page 2. This statement not consistent with the understanding of one of skill in the art of Pfam alignments. As known to those of skill in the art (and described in the Pfam documentation available at http://pfam.wustl.edu/faq.shtml), Pfam alignments do not display homology between pairs of sequences but rather display the fit of a particular query sequence to a particular protein family model. Thus the measure of the strength of a match between the query sequence and the Pfam consensus alignment is the Pfam bit score, which shows the statistical significance of the fit between the query sequence and the Pfam consensus alignment.

A Pfam "bit score" represents the log base 2 of the ratio of the probability of the sequence, given the hypothesis that the sequence belongs to the protein family being modeled versus the probability of the sequence given the hypothesis that the sequence was generated according to a random background model. Thus, the bit score of 145 for the 21612 polypeptide when fit to the Pfam short chain dehydrogenase model means that the 21612 amino acid sequence is 2^{145} (4.46×10^{43}) times more likely to belong to the short chain dehydrogenase family than to contain the amino acid sequence shown in Appendix B by chance. Accordingly, the bit score obtained when the 21612 polypeptide is fit to the Pfam ADH short chain dehydrogenase model strongly supports the functional determination for the 21612 polypeptide described in the specification.

B. The evidence presented by the Examiner does not support the Examiner's argument that the 21612 polypeptide is not an alcohol dehydrogenase.

The Examiner argues the Applicant's assertion that the 21612 polypeptide functions as an alcohol dehydrogenase is not persuasive because the 21612 polypeptide does not share a high level of overall sequence identity with a known alcohol dehydrogenase. In support of this arguments, the Examiner states that in a search of the sequence databases to identify sequences sharing sequence identity with the 21612 polypeptides, the 21612 polypeptide shared 41.7 % sequence identity with a ribotol dehydrogenase from *C. elegans* (NCBI Accession No. T19954) and 12.7% sequence identity with a human alcohol dehydrogenase (NCBI Accession No. AA622988). As an initial matter, Applicant notes that the function of the 21612 polypeptide was not determined based on overall sequence identity with the proteins, but rather based on Pfam analysis as described above. As described above, the, the bit score of 145 for the 21612 polypeptide when fit to the Pfam short chain dehydrogenase model means that the 21612 amino acid sequence is 2^{145} times more likely to belong to the short chain dehydrogenase family than to contain the amino acid sequence shown in Appendix B by chance.

Furthermore, with regards to NCBI Accession No. AA622988, Applicant notes that

nucleotides 27-386 of this EST sequence (which is only 386 nucleotides in length) share 99% sequence identity with nucleotides 1694-2052 of the nucleotide sequence set forth in SEQ ID NO:7. Accordingly, the level of sequence identity shared between NCBI Accession No. AA62298 and SEQ ID NO:7 supports rather than undermines the conclusion that the 21612 polypeptide has dehydrogenase activity.

In addition, although the polypeptide of NCBI Accession No. T19954 shares 41.7% overall sequence with the 21612 amino acid sequence shown in SEQ ID NO:7, it shares approximately 64% local sequence identity and approximately 74% local sequence similarity over amino acids 2 –276 of SEQ ID NO:7. This region of the 21612 polypeptide encompasses the Pfam short chain dehydrogenase consensus sequence shown in Appendix B. Those of skill in the art recognize that a high level of sequence identity within a functional domain of a polypeptide is a reliable indicator of polypeptide function. Therefore, the fact that the 21612 polypeptide and the dehydrogenase of NCBI Accession No. T19954 share 74% sequence similarity within a region encompassing the 21612 dehydrogenase domain supports the conclusion that the 21612 polypeptide functions as a dehydrogenase.

The sequence similarity shared between the 21612 polypeptide and NCBI Accession No. T19954 has even greater significance because of the evolutionary divergence between *Homo sapiens*, the source of the 21612 polypeptide, and *C. elegans*, the source of the polypeptide of NCBI Accession No. T19954. Those of skill in the art recognize that lower degrees of sequence identity may serve as a an indicator of function when the sequences that are aligned come from organisms that are as evolutionarily divergent as *C. elegans* and *Homo sapiens*. For example, the Δ9 fatty acid desaturase encoded by the *Saccharomyces cerevisiae* OLE1 gene can be functionally replaced by the rat stearoly-CoA desaturase gene despite the fact that the two proteins share only 36% overall identity. *See*, Stukey *et al.* 1990, *J. Biol. Chem.* 265:20144-20149, a copy of which is provided as Appendix C. Similarly, the polypeptides encoded by the human *PIG-B* gene and the *S. cerevisiae GPI10* gene both function in glycosyl-phosphatidylinositol anchor biosynthesis despite the fact that these polypeptides share only 33 %

overall identity. *See*, Sütterlin *et al.* (1998) *Biochem J.* 332:153-159, which is provided as Appendix D. As another example, the human and *S. cerevisiae* Cdc7 polypeptides have common biochemical and functional activities despite the fact that they share only 32% sequence identity. *See*, Sato *et al.* (1997) *EMBO J.* 16:4340-4351, which provided as Appendix E.

C. The function of the 21612 dehydrogenase was determined using methods that are accepted by those of skill in the art.

As described in Applicant's Amendments mailed November 19, 2001 and June 21, 2002, the use of sequence similarity, particularly sequence similarity within functional domains of proteins, is an art-accepted, reliable tool for predicting protein function. In fact, the art is replete with examples of proteins whose functions have been successfully determined based on sequence similarity within functional domains.

One example is seen in Nguyen *et al.* (2001) *Mol. Pharmacol.* 59:427-433, provided herewith as Appendix F. This reference describes the identification of the H4 receptor based on a query of GenBank to identify sequences sharing sequence similarity with GPCRs (see page 428, column 1, first full paragraph). Based on sequence similarity with the histamine receptor H3, the H4 polypeptide was predicted to have histamine receptor function. Subsequently, biochemical assays confirmed that histamine serves as a ligand for the H4 receptor and causes the receptor's internalization.

Another example is described in Dickman (1997) *Science* 277:1605-1606, a copy of which is provided as Appendix G. This reference describes the cloning and characterization of p73, a homolog of the tumor suppressor protein p53. The p73 polypeptide was predicted to have tumor suppressor activity based on its sequence similarity with p53 in several key domains, including the transcription activation domain, DNA binding domain, and oliogmerization domain (see page 160, column 3, first full paragraph). In accordance with this, p73 has subsequently been shown to have many of the activities of p53, including the ability to initiate cell death using the same pathway utilized by p53. Accordingly, sequence similarity to p53 was shown to be an

accurate predictor of p73 activity.

Yet another example is described in Kliewer *et al.* (1998) *Cell* 92:73-82, a copy of which is provided as Appendix H. This reference describes the identification of two novel nuclear receptors based on sequence similarity with the ligand binding domains of known nuclear receptors (see page 74, column 1, first full paragraph). The novel receptors, termed pregnane X receptor 1 and 2 (PXR.1 and PXR.2) were predicted to have nuclear receptor activity, *i.e.* hormone-regulated transcriptional activity, based on this sequence similarity. Biochemical evidence presented in the reference indicates that PXR regulates the transcription of the CYP3A family of steroid hydroxylases in a pregnane-dependent manner, confirming the functional predictions based on DNA binding domain and ligand binding domain sequence similarity.

The examples given here, which demonstrate the accuracy of sequence similarity-based predictions of protein function, represent only a few of the many such instances that are found in the scientific literature. Accordingly, one of skill in the art would find credible the Applicant's statement that the 21612 polypeptide functions as a dehydrogenase based on the Pfam analysis described above.

D. A knowledge of the tertiary structure of the 21612 dehydrogenase is not required in order to determine the function of this protein.

The Examiner argues that experimental evidence is required to demonstrate the function of the 21612 polypeptide because "it is general knowledge in the art that even conservative amino acid substitutions can adversely affect proper folding and biological activity if amino acids that art critical for such functions are substituted, and the relationship between the sequence of a polypeptide and its tertiary structure is neither well understood nor predictable . . .the ordinary artisan would immediately recognize that the encoded polypeptide must assume the proper three-dimensional configuration to be active, which is dependent upon the surrounding residues." August 14, 2002 Advisory Action, page 2. Applicant notes that the function of the 21612 dehydrogenase was not determined based on the predicted three-dimensional structure of this

polypeptide. Rather, the function of this protein was determined based on the high score that resulted when the 21612 sequence was fit with the Pfam model for ADH short chain dehydrogenases as described above. Accordingly, the Examiner's arguments are inapplicable to the methods used to determine 21612 function.

Furthermore, while Applicant agrees that certain regions of a protein must maintain a particular tertiary structure in order to retain activity, it does not follow that the structure of a protein must be known in order to know the function of the protein. In fact, although biochemical functions have been determined for thousands of proteins, three-dimensional structures have been determined for only a very small percentage of these biochemically-characterized proteins. Accordingly, the knowledge of a protein's tertiary structure is not a prerequisite to the determination of the protein's function. Therefore, the teaching that the prediction of protein tertiary structure based on primary sequence is difficult does not support the conclusion that the methods used by the Applicant to determine the function of the 21612 polypeptide are inaccurate.

E. Sequence Homology is an Acceptable Basis for Determining Utility According to United States Patent and Trademark Office Utility Examination Guidelines

The United States Patent and Trademark Office "Utility Examination Guidelines, 66 Fed Reg. 1092 (2001), make it clear that sequence homology is sufficient to establish utility, and that, contrary to statements made by the Examiner, biochemical evidence is not a *per se* requirement for the establishing the utility of a novel protein. The "Utility Examination Guidelines" state, "[w]hen a patent application claiming a nucleic acid asserts a specific, substantial, and credible utility, and bases the assertion upon homology to existing nucleic acids or proteins having an accepted utility, the asserted utility must be accepted by the examiner unless the Office has sufficient evidence or sound scientific reasoning to rebut such an assertion." 66 Fed. Reg. at 1096.

This statement in the "Utility Examination Guidelines" is illustrated in Example 10 of the "Revised Interim Utility Guidelines Training Material Examples, available at http://www.uspto.gov/web/offices/pac/utility/utilityguide.pdf. This example is directed to a polypeptide that shares sequence similarity with a DNA ligase. According to the analysis in the Example, "DNA ligases have a well-established use in the molecular biology art based on this class of protein's ability to ligate DNA. Revised Interim Utility Guidelines Training Materials, March 1, 2000, at 54. As in Example 10, the 21612 polypeptide has been shown to share sequence similarity with a protein family having a known biochemical activity. Accordingly, based on analogy to Example 10, the present invention also meets the criteria for well-established utility.

The sequence of Example 10 of the "Training Materials" is accorded to have a specific and substantial utility according to the "Revised Interim Utility Guidelines Training Material Examples," *ibid.*, despite the fact that the encoded polypeptide has not been directly demonstrated to have DNA ligase activity, and the substrate (i.e. single-stranded DNA or double-stranded DNA; blunt-ended DNA, 5' recessed ended DNA, 3' recessed ended DNA), co-factor requirements, and reaction conditions required to practice the invention are not disclosed. Thus, in accordance with the Utility Examination Guidelines, the very fact that the sequence of Example 10 has sequence similarity with a known protein possessing well-established utility is sufficient to confer a specific, substantial, and credible utility upon the claimed sequence. As the claimed invention of the present application is analogous to the situation described in Example 10, the criteria for utility have been met for claims 63-67, 77-79, and 87-104.

II. A prima facie showing of no utility has not been presented.

The "Examination Guidelines for the Utility Requirement" (MPEP § 2107) set forth the elements required to establish a *prima facie* case of no utility as follows:

> Where the asserted utility is not specific or substantial, a *prima facie* showing must establish that it is more likely than not that a person of ordinary skill in the art would not consider that any utility asserted by the applicant would

be specific and substantial. The prima facie showing must contain the following elements:

(i) An explanation that clearly sets forth the reasoning used in concluding that the asserted utility for the claimed invention is not both specific and substantial nor well-established;

(ii) Support for factual findings relied upon in reaching this conclusion; and

(iii) An evaluation of all relevant evidence of record, including utilities taught in the closest prior art.

MPEP § 2107 (8th ed. 2001). These guidelines are in accordance with *In re Brana*, 34 U.S.P.Q.2d 1437 (Fed. Cir. 1995), where the Federal Circuit held that, "[o]nly after the PTO provides evidence showing that one of ordinary skill in the art would reasonably doubt the asserted utility does the burden shift to the Applicant to provide rebuttal evidence sufficient to convince such a person of the invention's asserted utility." 34 U.S.P.Q.2d at 1441.

In the present case, the Applicant has use methods accepted by those of skill in the art to demonstrate that the 21612 polypeptide functions as an alcohol dehydrogenase, a class of polypeptides having well-established utility. Nevertheless, the Examiner has maintained the rejection under 35 U.S.C. § 101, stating "[t]he specification fails to show a single working example that establishes that the SEQ ID NO:8 which encodes the amino acid sequence of SEQ ID NO:7 is a member of Alcohol dehydrogenase (ADH) family." March 25, 2002 Office Action, page 3, and August 14, 2002 Advisory Action, page 2. In order to support this rejection in accordance with the "Examination Guidelines for the Utility Requirement," the Examiner must provide an explanation and factual findings showing that a person of ordinary skill in the art would not find Applicant's assertion that the 21612 polypeptide functions as an alcohol dehydrogenase to be credible. However, contrary to the requirements of MPEP § 2107, the Examiner has not provided factual findings or evidence that support this conclusion.

Furthermore, the Applicant has provided evidence demonstrating that those of skill in the art accept the methods used to determine the function of the 21612 polypeptide. Under the Guidelines, "[t]he examiner's decision [with respect to patentable utility] must be supported by a preponderance of all the evidence of record," MPEP § 2107.02 (8th ed. 2001), *citing In re Oetiker*, 24 U.S.P.Q.2d 1443 (Fed. Cir. 1992). In the present case, the Applicant's references support the finding that the 21612 polypeptide functions as an alcohol dehydrogenase, and the evidence presented by the Examiner would not lead one of skill in the art to question the credibility of this assertion. Accordingly, the preponderance of the evidence supports the conclusion that the 2162 polypeptide has patentable utility.

III. The invention of claims 63-67, 77-79, and 87-104 has utility under 35 U.S.C. §101.

The Applicant has demonstrated that the 21612 protein is a member of the alcohol dehydrogenase family, a protein family having well-established utility. The Applicant has further shown that the methods used to determine the function of the 21612 dehydrogenase are reliable and well-accepted by those of skill in the art. The Examiner has argued that Applicant's functional assignment for the 21612 polypeptide is not credible, but has not provided sufficient evidence or reasoning to establish a *prima facie* case of no utility. For these reasons, the rejection of claims 63-67, 77-79, and 87-104 under 35 U.S.C. §101 for lack of patentable utility should be reversed.

Issue 2--Whether the invention of claims 63-67, 77-79, and 87-104 is enabled under 35 U.S.C. § 112, first paragraph.

I. The Applicant has demonstrated that the 21612 polypeptide functions as an alcohol dehydrogenase.

The Examiner has maintained the rejection of claims 63-67, 77-79, and 87-104 under 35 U.S.C. § 112, first paragraph, on the grounds that the specification does not provide sufficient

guidance to enable one skilled in the art to make and use the claimed invention. The Examiner argues that the claimed invention is not enabled because "[i]t is unclear how one skill in the art would use the invention as claimed when the function of the polypeptide encoded by the nucleotide sequence of SEQ ID NO:8 is not known." March 25, 2002 Office Action, page 5. However, Applicant has shown that the 21612 polypeptide has dehydrogenase activity as described above in the arguments addressing the rejection under 35 U.S.C. § 101. Furthermore, the Examiner has presented no evidence to demonstrate that one of skill the art would doubt the credibility of Applicant's assertion that the 21612 polypeptide functions as a dehydrogenase. Accordingly, the premise on which this grounds for the rejection is based, i.e. that the 21612 polypeptide does not have dehydrogenase activity, is not supported by the evidence of record.

II. The specification provides sufficient guidance to allow one of skill in the art to make and use functional variants of the 21612 dehydrogenase.

The Examiner argues that the specification does not provide sufficient guidance to allow one of skill in the art to make and use variants of the 21612 polypeptide having dehydrogenase activity. In fact, sufficient guidance for making and using the claimed variant sequences is given in the specification. Applicant has provided the entire 21612 amino acid sequence, SEQ ID NO:7, as well as a nucleotide sequence which encodes it, SEQ ID NO:8 The variant nucleotide sequences of claims 88-100 vary from the nucleotide sequence of SEQ ID NO 8 by structural parameters (i.e. percent sequence identity to SEQ ID NO:8, hybridization with the complement of SEQ ID NO:8 under stringent conditions, or deletion) that are defined in the specification, and the claimed variants retain the dehydrogenase activity of the 21612 polypeptide having the amino acid sequence set forth in SEQ ID NO:7. Guidance for determining percent sequence identity and hybridization under stringent conditions is provided in the specification. *See*, page 18, line 4 *et seq.* and page 60, line 8 *et seq.* Polypeptide sequence variants that retain function are also described in the specification as containing "only conservative variation or variation in non-critical residues or in non-critical regions." *See*, page 21, lines 20-21 of the specification.

Guidance regarding conservative substitutions of amino acids is found in the specification on page 17, lines 6-12 and in Table 1.

Furthermore, the 21612 polypeptide shares a high level of sequence identity with a consensus domain that is conserved among members of the short chain dehydrogenase family. *See*, Appendix B. The specification also teaches methods for determining additional residues that are essential for function, including site-directed mutagenesis and alanine-scanning mutagenesis. *See*, page 22, lines 1-10.

Finally, the specification provides guidance regarding assays for dehydrogenase activity on page 22, lines 11-19. Accordingly, one of skill in the art would be able to determine the functionality of 21612 variants.

Thus, a rational scheme for determining the regions of the 21612 short chain dehydrogenase that would tolerate modification is provided. Based on the regions of the 21612 polypeptide that are conserved with other short chain dehydrogenases, and the methods provided for identifying additional residues critical for 21612 function, the skilled artisan could choose among possible modifications to produce polypeptides encodes by nucleotide sequences within structural parameters set forth in the claims and then test these modified variants to determine if they retain dehydrogenase activity. Although some quantity of experimentation would be required, the level of experimentation would not be undue in view of the amount of direction provided in the specification, the state of the prior art, and the level of skill of one of ordinary skill in the art. These factors all favor a conclusion that one of skill in the art could practice the claimed invention without undue experimentation.

Applicant notes that an enabling disclosure need describe the claimed invention in such a way as to enable the ordinarily skilled artisan to make and use the invention, and that this description be commensurate with the scope of the claimed invention. There is no requirement that the disclosure provide working examples of every permutation of the invention.

Nevertheless, the Examiner has maintained the rejection on the grounds that "the specification fails to disclose that any and all variants of SEQ ID NO:7 (as claimed) are capable

of eliciting any ADH-like activity." March 25, 2002 Office Action, page 5. However, Applicant has **not** asserted that every nucleotide sequence meeting the structural limitations of claims 88-92 will also encode a polypeptide having dehydrogenase activity. Rather, these claims are limited to nucleotide sequences meeting both the structural requirements of these claims (i.e. the nucleotide sequences share sequence identity with the nucleotide sequence set forth in SEQ ID NO:8, hybridize with the nucleotide sequence set forth in SEQ ID NO:8 under specified conditions, or contain a subset of the nucleotide sequence set forth in SEQ ID NO:8 or the nucleotide sequence of the cDNA insert of the plasmid deposited with ATCC as patent deposit number PTA-2170) and the functional requirements of these claims (i.e. the nucleotide sequences encode a polypeptide having dehydrogenase activity). Furthermore, as described in Applicant's response mailed November 19, 2001, the specification provides sufficient guidance to allow one of skill in the art to make nucleotide sequences falling within the structural limitations of the claims and determine whether these sequences encode polypeptides having the functional limitation of the claims. Accordingly, the scope of enablement provided in the specification is commensurate with the scope of the claims.

Applicant cites Gayle *et al.* (1993) *J. Biol. Chem.* 268:22105-22111, provided herewith as Appendix I to demonstrate that making functional variants of polypeptides is routine to those of skill in the art and does not require undue experimentation. This reference describes saturation mutagenesis of the mature human interluekin-1α (IL-1α) sequence. The authors report that more than 3,500 mutants of IL-1α were examined to determine their biological activity and their binding activity, demonstrating that such mutants can be made and tested for function without undue experimentation. Furthermore, the authors state that "[m]ost of the molecule could be mutated with little effect on either [biological or binding] activity." Gayle *et al.* at 22109, column 2.

The Examiner has argued that the specification does not enable variants of the 21612 polypeptide on the grounds that "[i]t is general knowledge in the art that even conservative amino acid substitutions can adversely affect proper folding and biological activity if amino acids that

are critical for such functions are substituted, and the relationship between the sequence of the polypeptide and its tertiary structure is neither well understood nor predictable." March 25, 2002 Office Action, page 6. The premise of the Examiner's argument appears to be that Applicant must demonstrate a working example of each and every sequence falling within the limitation of the claims, and that the claimed invention is enabled only if no experimentation is required to make and use the claimed variants. This requirement is not supported by the applicable case law. The test of enablement is not whether experimentation is necessary, but rather if experimentation *is* necessary, whether it is undue. *In re Angstadt*, 198 USPQ 214, 219 (C.C.P.A. 1976). Factors to be considered in determining whether undue experimentation is required include the quantity of experimentation necessary, the amount of guidance provided in the specification, the presence of working examples of the invention in the application, the nature of the invention, the state of the prior art, the relative skill of those in the art, the predictability in the art, and the breadth of the claimed invention. *In re Wands*, 8 USPQ 2d 1400, 1404 (Fed. Cir. 1988).

Based on the guidance provide in the specification, the regions of the 21612 polypeptide that are conserved with other short chain dehydrogenases, and the methods provided for identifying additional residues critical for 21612 function, the skilled artisan could choose among possible modifications to produce polypeptides encoded by nucleotide sequences within structural parameters set forth in the claims and then test these modified variants to determine if they retain dehydrogenase activity. Although some quantity of experimentation would be required, the level of experimentation would not be undue in view of the quantity of experimentation necessary, the amount of direction provided in the specification, the state of the prior art, the presence of a working example of the invention, the level of skill of one of ordinary skill in the art, and the breadth of the claimed invention. These factors all favor a conclusion that one of skill in the art could practice the claimed invention without undue experimentation.

III. The case law cited by the Examiner does not support the rejection for lack of enablement under 35 U.S.C. §, 112, first paragraph.

In the Advisory Action mailed August 14, 2002, the Examiner cites *In re Fisher*, 166 USPQ 19 (CCPA 1970), *Hybritech Inc. V. Monoclonal Antibodies, Inc.* 231 USPQ 81 (Fed. Cir. 1986), *Genentech Inc. v. Novo Nordisk* 42 USPQ2d 1005 (CAFC 1997), *In re Wands* 8 USPQ2d 1400 (Fed. Cir. 1988) and *Ex parte Singh* 17 USPQ2d (BPAI 1991) in support of the argument that Applicant must teach each and every sequence that may that may be added, deleted, or substituted to the sequences specifically disclosed. However, none of the cited cases support the enablement standard applied by the Examiner.

The facts of *Fisher* are distinguishable from the present case, because in *Fisher* the applicants had claimed adrenocorticotrophic hormones (ACTHs) of "at least 24 amino acids." The Court found that, at the time (the 1960s), the art suggested that all known ACTHs contained 39 amino acids and that one of skill in the art would not know how to make other ACTHs. In the present case, by contrast, guidance regarding the conserved regions of dehydrogenase is both known in the art and provided in the specification. Moreover, methods for making altered sequences are routine in the art, as are techniques for testing for alcohol dehydrogenase activity. Thus, *Fisher* is inapposite to the present application.

In *Hybritech*, the issue before the court was whether the specification of U.S. Patent No. 4,376,110 failed to provide sufficient enablement for the claimed methods of determining antibody concentration because the specification did not disclose how to make monoclonal antibodies, screen for proper monoclonal antibodies, and measure monoclonal antibody affinity. The court held that enablement "is not precluded even if some experimentation is necessary, although the amount of experimentation must not be unduly extensive . . .[f]urthermore, a patent need not teach, and preferably omits, what is well known in the art." 231 USPQ at 94. The court found that, based on the evidence presented at trial, "there was not a shred of evidence that undue experimentation was required by those skilled in the art to practice the invention." *Id.* Similarly, in the present application, methods of making and testing variant polypeptides are well known in

the art and thus undue experimentation is not required to make and use the claimed 21612 variants.

In *Wands*, the court held that the test of whether an invention requires undue experimentation is not based on a single factor, but rather a conclusion reached by weighing many factors. The factors to be considered in determining whether undue experimentation is required include the quantity of experimentation necessary, the amount of guidance provided in the specification, the presence of working examples of the invention in the application, the nature of the invention, the state of the prior art, the relative skill of those in the art, the predictability in the art, and the breadth of the claimed invention. 8 USPQ2d at 1404.

Accordingly, the holding of *Wands* does not require that an applicant identify each and every operable embodiment as asserted by the Examiner. Rather, *Wands* sets out factors to be considered in determining whether undue experimentation is required to practice the claimed invention. The Applicant has provided the novel polypeptide of SEQ ID NO:7 and has demonstrated that this protein is a member of the ADH short chain dehydrogenase family. Thus, contrary to statements made in the office action, a working example of the claimed invention is provided. Furthermore, based on the guidance provide in the specification, the regions of the 21612 polypeptide that are conserved with other ADH family dehydrogenases, and the methods provided for identifying additional residues critical for 21612 function, the skilled artisan could choose among possible modifications to produce polypeptides encoded by nucleotide sequences within structural parameters set forth in the claims and then test these modified variants to determine if they retain dehydrogenase activity. Although some quantity of experimentation would be required, the level of experimentation would not be undue in view of the quantity of experimentation necessary, the amount of direction provided in the specification, the state of the prior art (as demonstrated by Gayle *et al.*), the presence of a working example of the invention, the level of skill of one of ordinary skill in the art, for whom the making and testing of alcohol dehydrogenase variants is routine. These factors all favor a conclusion that one of skill in the art could practice the claimed invention without undue experimentation.

The Examiner cites *Genentech* for the court's statement that "[i]t is the specification, not the knowledge of one skilled in the art, that must supply the novel aspects of an invention in order to constitute enablement." 42 USPQ2d at 1005. However, in the present case the specification does provide the novel aspect of the invention, i.e., the 21612 dehydrogenase sequences given in SEQ ID NO:7 and SEQ ID NO:8. Once these novel sequences are disclosed, it is routine to those of skill in the art to make functional variants as illustrated by the Gayle *et al.* reference cited above. In *Genentech*, the court held that "where there is no disclosure of any specific starting material or of any of the conditions under which a process can be carried out, undue experimentation is required." 42 USPQ2d at 1005. In the present case, the specification provides both the starting material, *i.e.* the 21612 sequence, and a rational scheme for making functional variants and fragments. Accordingly, the holding of *Genentech* is inapposite to the facts of the present case.

In *Singh*, the issue before the Board of Patent Appeals and Interferences was whether the specification of U.S. Patent Application No. 06/506,098 provided sufficient guidance to enable claims to a process of expressing heterologous polypeptides in yeast. The Board found that the claims were not sufficient enabled because the complete secretory pathway of yeast was not known at the time the invention was made and thus is was not predictable that any protein other than the one specifically disclosed in the specification could be successfully expressed in yeast. The Examiner cites this case for the court's statement that "[i]t has been recognized that the unpredictability of a particular art area may alone provide a reasonable doubt as to the accuracy of a broad statement made in support of the enablement of a claim." 17 USPQ2d at 1715. However, the Examiner has presented no evidence to demonstrate that making and using functional variants of the 21612 polypeptide is sufficiently unpredictable to constitute undue experimentation. Furthermore, the Gayle *et al.* reference provided by Applicant as Appendix I demonstrates that making and using functional variants of a known sequence is routine to those of skill in the art. Accordingly, the holding of *Singh* is inapposite to the facts of the present case.

IV. The specification provides enablement for the host cells of claims 65-67 and 95-97 that is commensurate with the scope of these claims.

The Examiner has rejected claims 65-67 and 95-97 on the grounds that these claims encompass host cells *in vivo* and therefore Applicant must provide guidance for the use of the sequences in gene therapy and in producing transgenic animals. The Examiner argues that "gene therapy is considered a highly experimental area of research at this time, and both researchers and the public agree that demonstratable progress to date has fallen short of initial expectations." March 25, 2002 Office Action, page 7. The Examiner also argues that "the phenotype of an animal is determined by a complex interaction of genetics and environment." March 25, 2002, Office action, page 7. However, the arguments presented in the Office Action are not relevant to the claimed subject matter because the claims are not directed to methods of gene therapy or to methods of producing a transgenic animal having a particular phenotype but instead are directed to host cells containing specified nucleic acid molecules. Methods of producing such host cells are described on pages 85-90 of the specification. Based on the guidance provided, one of skill in the art could produce the claimed host cells without undue experimentation. The proper test for the enablement of an invention is whether the specification provides enablement commensurate with the scope of what is claimed. The specification has provided sufficient guidance to allow one of skill in the art to make and use the host cells of claims 65-67 and 95-97, and therefore the enablement requirement is met.

V. The invention of claims 63-67, 77-79, and 87-104 is enabled under 35 U.S.C. § 112, first paragraph.

In *Genentech v. Novo Nordisk*, 42 USPQ 1001 (Fed. Cir. 1997) the Federal Circuit held that "[t]o be enabling, the specification of a patent must teach those skilled in the art how to make and use the full scope of the claimed invention **without 'undue experimentation.'** " 42 USPQ at 1004, citing In re Wright, 999 F.2d 1557, 1561, 27 USPQ2d 1510, 1513 (Fed.Cir.1993); Amgen, Inc. v. Chugai Pharm. Co., 927 F.2d 1200, 1212, 18 USPQ2d 1016,

1026 (Fed.Cir.1991); In re Fisher, 57 C.C.P.A. 1099, 427 F.2d 833, 839, 166 USPQ 18, 24 (1970), *emphasis added.* Applicant has provided the sequence of a novel dehydrogenase, provided a rational scheme for making functional variants and fragments of this dehydrogenase, and demonstrated that making such variants and fragments is routine to those of skill in the art. Accordingly, the requirements for enablement have been met, and the rejection of claims 63-67, 77-79, and 87-104 under 35 U.S.C. §112, first paragraph, for lack of enablement should be reversed.

Issue 3--Whether a sufficient written description of the invention of claims 88-92 is provided under 35 U.S.C. § 112, first paragraph.

I. The Applicant has demonstrated that the 21612 polypeptide functions as an alcohol dehydrogenase.

The Examiner has maintained the rejection of claims 88-92 under 35 U.S.C. § 112, first paragraph, on the grounds that the specification does not provide an adequate written description for the claimed invention because "the applicant fails to point out where in the specification it is disclosed that the polypeptide encoded by the nucleic acid molecule of SEQ ID NO:8 have any alcohol-dehydrogenase-like activity explicitly or implicitly as putatively consider by the applicant." March 25, 2002 Office Action, page 8. However, Applicant has shown that the 21612 polypeptide has dehydrogenase activity as described in the specification and addressed in the arguments above relating to the rejection under 35 U.S.C. § 101. Furthermore, the Examiner has presented no evidence to demonstrate that one of skill the art would doubt the credibility of Applicant's assertion that the 21612 polypeptide functions as a dehydrogenase. Accordingly, the premise on which the rejection is based, i.e. that the 21612 polypeptide does not have dehydrogenase activity, is not supported by the evidence of record.

II. An improper standard has been applied in determining whether the specification provides a sufficient written description of the invention of claims 88-92.

The Examiner states that the rejection under 35 U.S.C. § 112, first paragraph for lack of written description in maintained because "[t]he specification fails to disclose any and all variants of nucleic and amino acid sequences of SEQ ID NO(s) as claimed." March 25, 2002 Office Action, page 8. This statement suggests that the Applicant must disclose the sequence of each variant falling within the structural and functional limitations set forth in the claims in order to adequately describe the claimed genus of sequences. However, the requirement set forth in the office action is not supported by the "Guidelines for Examination of Patent Applications Under the 35 U.S.C. §112, ¶ 1, 'Written Description' Requirement" (66 Fed. Reg. 1099 (2001)) or the supporting case law.

The "Guidelines for Examination of Patent Applications Under the 35 U.S.C. 112, ¶ 1, 'Written Description' Requirement" state that genus may be described by "sufficient description of a representative number of species . . . or by disclosure of relevant, identifying characteristics, *i.e.* structure or other physical and/or chemical properties." *Id.* at 1106. This requirement is in accordance with *Regents of the University of California v. Eli Lilly & Co.*, 119 F.3d 1559 (Fed. Cir. 1997), where the court held that "[a] written description of an invention involving a chemical genus, like a description of a chemical species, 'requires a precise definition, such as by structure, formula, or chemical name' of the claimed subject matter sufficient to distinguish it from other materials." 119 F.3d at 1568, citing Fiers v. Revel 984 F.2d 1164 (Fed. Cir. 1993).

Applicant submits that the written description provided for the sequences recited in claims 88-92 meets this requirement. These claims recite the identifying structural characteristics that define each genus of nucleotide sequences. Claims 88-90 recite nucleotide sequences having at least 70%, 80%, or 90% sequence identity with the nucleotide sequence set forth in SEQ ID NO:8, claim 91 recites nucleotide sequences that hybridize to the nucleotide sequence set forth in SEQ ID NO:8 under specified conditions, and claim 92 recites nucleotide sequences encoding a fragment of the amino acid sequence set forth in SEQ ID NO:7 or the

amino acids sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Patent Deposit Number PTA-2170, wherein the fragment has dehydrogenase activity and consists of at least 139 contiguous amino acids of the amino acid sequence set forth in SEQ ID NO:7 or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Patent Deposit Number PTA-2170. These structural limitations are sufficient to distinguish the claimed nucleotide sequences from other materials and thus sufficiently define the claimed genus.

Furthermore, in *Regents of the University of California v. Eli Lilly & Co*, 119 F.3d 1559 (Fed. Cir. 1997), the court held that "[a] description of a genus of cDNAs may be achieved by means of a recitation of a representative number of cDNAs, defined by nucleotide sequence, falling within the scope of the genus or of a recitation of structural features common to the members of the genus, which features constitute a substantial portion of the genus." 119 F.3d at 1569. The recitation of the structural features of sequence identity with SEQ ID NO:7, hybridization with SEQ ID NO:7, or the presence of subsequences of SEQ ID NO:8 or the amino acid sequence encoded by the plasmid deposited with ATCC as Patent Deposit Number PTA-2170, where the fragments have a given minimum length, is sufficient to satisfy this requirement.

Applicant has further provided the functional characteristics that distinguish the claimed sequences of the genus. Specifically, claims 88-92 recite that the variants and fragments have dehydrogenase activity. Accordingly, both the structural properties and the functional properties that characterize the claimed genus are specifically recited in the claims.

The present claims are analogous to those presented in Example 14 of the "Revised Interim Written Description Guidelines Training Materials," available at www.uspto.gov/web/menu/written.pdf. Example 14 is directed to a protein having at least 95% sequence identity to the sequence of SEQ ID NO:3, wherein the sequence catalyzes the reaction A → B. The conclusion in the Training Materials is that the generic claim of Example 14 is sufficiently described under § 112, first paragraph, because 1) "the single sequence disclosed in SEQ ID NO:3 is representative of the genus" and 2) the claim recites a limitation requiring the

compound to catalyze the reaction from A → B. Therefore one of skill in art would recognize that the Applicant was in possession of the necessary common attributes possessed by the members of the genus.

Following the analysis of Example 14, Applicant submits that claims 88-92 satisfy the written description requirements of § 112, first paragraph. Specifically, the claims of the present invention encompass nucleotide sequences having sequence identity to the nucleotide sequence of SEQ ID NO:8, hybridizing under stringent conditions to the nucleotide sequence of SEQ ID NO:8, comprising a subsequence of SEQ ID NO:7, wherein the claimed sequences encode a polypeptide having a specified activity. As in Example 14, the specification discloses the nucleic acid sequence of SEQ ID NO:1, and the claims recite a limitation requiring the compound to have a specific function (*i.e.* dehydrogenase activity). Accordingly, claims 88-92 provide the relevant, identifying characteristics that describe the claimed genus, and one of skill in the art would recognize that the inventors were in possession of the claimed invention.

III. The case law cited by the Examiner does not support the rejection for insufficient written description under 35 U.S.C. §, 112, first paragraph.

The Examiner cites *In re Shokal* 113 USPQ 283 (CCPA 1957), *Purdue Pharma L.P. v. Faulding Inc.*56 USPQ2d 1481 (Fed. Cir. 2000), and *Fiddes v. Baird*, 30 USPQ2d 1481 (BPAI 1993) in support of the rejection for insufficient written description. However, none of the cited cases support the standard for written description applied by the Examiner.

The Examiner cites *Shokal* in support of the argument that a single species can rarely be used to support a generic claim. However, the facts of the present case are readily distinguishable from those in *Shokal*. In *Shokal*, the specification filed by the applicants did not define the claimed genus by identifying characteristics that would distinguish the claimed invention from other compounds. Instead, the applicants attempted to establish possession of the claimed genus by the disclosure of representative species.

In contrast, the claimed genus in the present application is defined in the specification by the functional properties (*i.e.* dehydrogenase activity) and the structural properties (*i.e.* sequence identity with disclosed sequences, hybridization with disclosed sequences, or deletions of disclosed sequences) shared by the encompassed species. This description is sufficient to distinguish the claimed genus from other materials. Accordingly, the disclosure of many species within the claimed genus is not required. Indeed, the court in *Shokal* makes this distinction, stating, "*where the genus is not set forth in express terms*, the number and nature of the examples given, together with the accompanying disclosure must be such as to indicate clearly what the genus actually is." 113 USPQ at 285, *emphasis added.*

The Examiner also cites *Purdue Pharma* in support of the argument that a single species is rarely sufficient to describe a broad genus. The issue before the court in *Purdue Pharma* was whether the specification of U.S. Patent No. 5,672,360 provided a sufficient written description for a claim to a method for treating pain using an opioid at concentration such that the maximum plasma concentration (C_{max}) was more than twice the plasma level of the opioid 24 hours after administration (C_{24}). The court found that "there is nothing in the written description . . . that would suggest to one skilled in the art that the C_{max}/C_{24} ratio is an important defining quality of the formulation." 56 USPQ2d at 1486. The court further stated that "[w]hat the '350 patentees have done is to pick a characteristic possessed by two of their formulations, a characteristic that is not discussed even in passing in the disclosure, and then make it the basis of claims that cover not just those two formulations, but any formulation that has that characteristic." USPQ2d at 1487. In contrast, the structural and functional limitations recited in claims 88-92 are specifically described in the specification. Accordingly, the present case is factually distinguishable from *Purdue Pharma*, and this case does not support the rejection of claims 88-92 for insufficient written description.

In *Fiddes*, the issue before the Board of Patent Appeals and Interferences was whether a count directed to a recombinant DNA molecule encoding any mammalian basic fibroblast growth factor (FGF) was sufficiently described by the disclosure of the amino acid sequence of bovine

FGF. The Board found that the claimed subject matter was not sufficiently described because the specification had described **no naturally-occurring genes** encoding FGF and thus had not described even one species falling within the claimed genus of DNA molecules. The Board stated that the inventor "was not in possession of the naturally occurring gene encoding bovine pituitary FGF . . . thus Baird was not in possession of the genes encoding mammalian FGFs since in 1987, knowledge of the amino acid sequence of a protein coupled with the established relationship between a nucleic acid and the protein it encodes would not establish possession of the gene encoding that protein." 30 USPQ2d at 1483. In contrast, the present application does provide the naturally-occurring human nucleotide sequence encoding the 21612 dehydrogenase. In addition, claims 88-92, unlike the count that was at issue in *Fiddes*, specifically recite the structural features that characterize the claimed genus of nucleotide sequences. As described above, these structural features are sufficient to distinguish the claimed nucleotide sequences from other materials and therefore satisfy the requirements of 35 U.S.C. § 112, first paragraph.

IV. The invention of claims 88-92 is sufficiently described under 35 U.S.C. § 112, first paragraph.

In *Regents of the University of California v. Eli Lilly & Co.*, 119 F.3d 1559 (Fed. Cir. 1997), the court held that "[a] written description of an invention involving a chemical genus, like a description of a chemical species, 'requires a precise definition, such as by structure, formula, or chemical name' of the claimed subject matter sufficient to distinguish it from other materials." 119 F.3d at 1568, citing Fiers v. Revel 984 F.2d 1164 (Fed. Cir. 1993). Claims 88-92 recite both the structural features and the functional properties that characterize the claimed genus of nucleotide sequences and distinguish them from other materials. Accordingly, claims 88-92 provide the relevant, identifying characteristics that describe the claimed genus, and one of skill in the art would recognize that the inventors were in possession of the claimed invention. For these reasons, the rejection of claims 88-92 under 35 U.S.C. §112, first paragraph, for insufficient written description should be reversed.

Issue 4--Whether claim 79 is indefinite under 35 U.S.C. § 112, second paragraph.

The Examiner has maintained the rejection of claim 79 under 35 U.S.C. § 112, second paragraph on the grounds that the claim is indefinite because "it is unclear what the 'instructions for use' would be in this context." August 14, 2002 Office Action, page 5. Claim 79 recites, "[a] kit for use in the method of claim 77, wherein said kit comprises at least one nucleic acid probe of claim 77 and instructions for use in the method of claim 77." Claim 77 recites, "[a] method for detecting the presence of a nucleic acid molecule of claim 87 in a sample, said method comprising the steps of contacting the sample with a nucleic acid probe which selectively hybridizes to the nucleic acid molecule and determining whether the nucleic acid probe binds to the nucleic acid molecule in the sample . . ." Thus, claim 77 specifically recites the essential steps of the claim method of detection, and one of skill in the art would recognize what is intended by the phrase "instructions for use in the method of claim 77." For these reasons, the rejection of claim 79 under 35 U.S.C. §112, second paragraph, should be reversed.

CONCLUSION

In view of the arguments presented above, Applicant contends that each of claims 63-67, 77-79, and 87-104 is patentable. Therefore, reversal of the rejections under 35 U.S.C. §101 and 35 U.S.C. §112, first and second paragraphs, is respectfully solicited.

Respectfully submitted,

Kathryn L. Coulter

Kathryn L. Coulter
Registration No. 45,889

CUSTOMER NO. 000826
ALSTON & BIRD LLP
Bank of America Plaza
101 South Tryon Street, Suite 4000
Charlotte, NC 28280-4000
Tel Raleigh Office (919) 862-2200
Fax Raleigh Office (919) 862-2260

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Box AF, Commissioner for Patents, Washington, DC 20231 on September 24, 2002.

Nora C. Martinez

Nora C. Martinez

In re: Meyers
Appl. No. 09/464,039
Filed December 15, 1999

APPENDIX A

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re:	Meyers	Confirmation No.:	7067
Appl. No.:	09/464,039	Group Art Unit:	1636
Filed:	December 15, 1999	Examiner:	Sumesh Kaushal
For:	21612, NOVEL HUMAN DEHYDROGENASE [TITLE AS AMENDED 11/19/01]		

APPEALED CLAIMS

63. The nucleic acid molecule of claim 87 further comprising vector nucleic acid sequences.

64. The nucleic acid molecule of claim 87 further comprising nucleic acid sequences encoding a heterologous polypeptide.

65. A host cell which contains the nucleic acid molecule of claim 87.

66. The host cell of claim 65 which is a mammalian host cell.

67. A nonhuman mammalian host cell containing the nucleic acid molecule of claim 87.

77. A method for detecting the presence of a nucleic acid molecule of claim 87 in a sample, said method comprising the steps of contacting the sample with a nucleic acid probe which selectively hybridizes to the nucleic acid molecule and determining whether the nucleic acid probe binds to the nucleic acid molecule in the sample; wherein said nucleic acid probe is selected from the group consisting of:

a) the nucleotide sequence set forth in SEQ ID NO:8;

b) the nucleotide sequence of a fragment of the nucleotide sequence set for in SEQ ID NO:8, wherein said fragment comprises at least 417 contiguous nucleotides of the nucleotide sequence set forth in SEQ ID NO:8;

c) a nucleotide sequence having at least 70% sequence identity to the nucleotide sequence set forth in SEQ ID NO:8; and

d) a nucleotide sequence complementary to a nucleotide sequence of a), b), or c).

78. The method of claim 77, wherein the sample comprises mRNA molecules.

79. A kit for use in the method of claim 77, wherein said kit comprises at least one nucleic acid probe of claim 77 and instructions for use in the method of claim 77.

87. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

a) the nucleotide sequence set forth in SEQ ID NO:8;

b) the nucleotide sequence of the cDNA insert of the plasmid deposited with ATCC as Patent Deposit Number PTA-2170;

c) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:7;

d) a nucleotide sequence encoding the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Patent Deposit Number PTA-2170; and

e) a nucleotide sequence complementary to a nucleotide sequence of a), b), c), or d).

88. An isolated nucleic acid molecule having a nucleotide selected from the group consisting of:

a) a nucleotide sequence encoding a polypeptide having dehydrogenase activity, wherein said nucleotide sequence has at least 70% sequence identity with the nucleotide sequence set forth in SEQ ID NO:8; and

b) a nucleotide sequence complementary to the nucleotide sequence of a).

89. The nucleic acid molecule of claim 88, wherein said nucleotide sequence is selected from the group consisting of:

a) a nucleotide sequence encoding a polypeptide having dehydrogenase activity, wherein said nucleotide sequence has at least 80% sequence identity with the nucleotide sequence set forth in SEQ ID NO:8; and

b) a nucleotide sequence complementary to the nucleotide sequence of a).

90. The nucleic acid molecule of claim 89, wherein said nucleotide sequence is selected from the group consisting of:

a) a nucleotide sequence encoding a polypeptide having dehydrogenase activity, wherein said nucleotide sequence has at least 90% sequence identity with the nucleotide sequence set forth in SEQ ID NO:8; and

b) a nucleotide sequence complementary to the nucleotide sequence of a).

91. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

a) a nucleotide sequence encoding a polypeptide having dehydrogenase activity, wherein the complement of said nucleotide sequence hybridizes under stringent conditions to the nucleotide sequence set forth in SEQ ID NO:8, said stringent conditions comprising hybridization at about 45°C, followed by at least one wash in 0.2X SSC/0.1% SDS at 65°C; and

b) a nucleotide sequence complementary to the nucleotide sequence of a).

92. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of

a) a nucleotide sequence encoding a fragment of the amino acid sequence set forth in SEQ ID NO:7, wherein said fragment has dehydrogenase activity and consists of at least 139 contiguous amino acids of SEQ ID NO:7; and

b) a nucleotide sequence encoding a fragment of the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Patent Deposit Number PTA-2170, wherein the fragment has dehydrogenase activity and consists of at least 139 contiguous amino acids of the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Patent Deposit Number PTA-2170; and

c) a nucleotide sequence complementary to the nucleotide sequence of a) or b).

93. The nucleic acid molecule of claim 88 further comprising vector nucleic acid sequences.

94. The nucleic acid molecule of claim 88 further comprising nucleic acid sequences encoding a heterologous polypeptide.

95. A host cell which contains the nucleic acid molecule of claim 88.

96 The host cell of claim 95, wherein said host cell is a mammalian host cell.

97. A nonhuman mammalian host cell containing the nucleic acid molecule of claim 88.

98. A method for detecting the presence of a nucleic acid molecule of claim 88 in a sample, said method comprising the steps of contacting the sample with a nucleic acid probe which selectively hybridizes to the nucleic acid molecule and determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample; wherein said nucleic acid probe is selected from the group consisting of:

a) the nucleotide sequence set forth in SEQ ID NO:8;

b) the nucleotide sequence of a fragment of the nucleotide sequence set for in SEQ ID NO:8, wherein said fragment comprises at least 417 contiguous nucleotides of the nucleotide sequence set forth in SEQ ID NO:8;

c) a nucleotide sequence having at least 70% sequence identity to the nucleotide sequence set forth in SEQ ID NO:8; and

d) a nucleotide sequence complementary to a nucleotide sequence of a), b), or c).

99. The method of claim 98, wherein the sample comprises mRNA molecules.

100. A kit for use in the method of claim 98, wherein said kit comprises at least one nucleic acid probe of claim 98 and instructions for use in the method of claim 98.

101. A method for producing a polypeptide, said method comprising culturing a host cell containing a nucleic acid molecule of claim 87 under conditions in which the polypeptide encoded by the nucleic acid molecule is expressed.

102. A method for producing a polypeptide, said method comprising culturing a host cell containing a nucleic acid molecule of claim 88 under conditions in which the polypeptide encoded by the nucleic acid molecule is expressed.

103. A method for producing a polypeptide, said method comprising culturing a host cell containing a nucleic acid molecule of claim 90 under conditions in which the polypeptide encoded by the nucleic acid molecule is expressed.

104. A method for producing a polypeptide, said method comprising culturing a host cell containing a nucleic acid molecule of claim 91 under conditions in which the polypeptide encoded by the nucleic acid molecule is expressed.

Protein Family / Domain Matches, HMMer version 2

```
Searching for complete domains
hmmpfam - search a single seq against HMM database
HMMER 2.1.1 (Dec 1998)
Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL).
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HMM file:                 /prod/ddm/seqanal/PFAM/pfam4.3/Pfam
Sequence file:            /prod/ddm/wspace/orfanal/oa-script.6378.seq
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Query:  21612

Scores for sequence family classification (score includes all domains):
Model          Description                                    Score    E-value  N
--------       -----------                                    -----    ------- ---
adh_short      short chain dehydrogenase                      145.0    1.3e-39   1
beta-lactamase Beta-lactamase                                   3.3        6.6   1

Parsed for domains:
Model          Domain  seq-f seq-t    hmm-f hmm-t       score  E-value
--------       ------- ----- -----    ----- -----       -----  -------
adh_short        1/1      11   204 ..     1   203 []    145.0  1.3e-39
beta-lactamase   1/1     222   236 ..   317   331 .]      3.3      6.6

Alignments of top-scoring domains:
adh_short: domain 1 of 1, from 11 to 204: score 145.0, E = 1.3e-39
                   *->KvaLvTGassGIGlaiAkrLakeGakVvvadrneeklekG.....av
                       ++ +TGas+GIG+aiA+  ak+Ga++v+a+ ++ +  k+ ++    +
         21612    11    CTVFITGASRGIGKAIALKAAKDGANIVIAAKTAQPHPKLlgtiyTA 57

                      akelkelGgndkdralaiqlDvtdeesv.aaveqaverlGrlDvLVNNAG
                      a+e+++ Gg     +al++ +Dv+de++++aave+a++++G++D+LVNNA
         21612    58 AEEIEAVGG----KALPCIVDVRDEQQIsAAVEKAIKKFGGIDILVNNAS 103

                      giillrpgpfaelsrtmeedwdrvidvNltgvfllravlplmamkkrgg
                         ++  ++ ++ +     +  d +++vN +g+ l ++a++p    +kk++
         21612   104 ---AISLTNTLDTP---TKRLDLMMNVNTRGTYLASKACIP--YLKKSKV 145

                      GrIvNiSSvaGrkeg.glvgvpgg..saYsASKaAvigltrsLAlElaph
                        I NiS        +  l++v++ +++aY+ +K ++ +    ++A E +
         21612   146 AHILNISP------PlNLNPVWFKqhCAYTIAKYGMSMYVLGMAEEFKGE 189

                      gIrVnavaP.GgvdTd<-*
                       I Vna  P+ ++ T+
         21612   190 -IAVNALWPkTAIHTA    204

beta-lactamase: domain 1 of 1, from 222 to 236: score 3.3, E = 6.6
                   *->eliaeaakvvleaLg<-*
                      ++ia+aa++++++ +
         21612   222    DIIADAAYSIFQKPK    236

//
```

APPENDIX B

In re: Meyers
Appl. No. 09/464,039
Filed December 15, 1999

APPENDIX C

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 265, No. 33, Issue of November 25, pp. 20144-20149, 1990
Printed in U.S.A.

The *OLE1* Gene of *Saccharomyces cerevisiae* Encodes the Δ9 Fatty Acid Desaturase and Can Be Functionally Replaced by the Rat Stearoyl-CoA Desaturase Gene*

(Received for publication, June 12, 1990)

Joseph E. Stukey‡, Virginia M. McDonough§, and Charles E. Martin¶

From the Nelson Biological Laboratory, Bureau of Biological Research, Rutgers University, Busch Campus, Piscataway, New Jersey 08855-1059

Strains of *Saccharomyces cerevisiae* bearing the *ole1* mutation are defective in unsaturated fatty acid (UFA) synthesis and require UFAs for growth. A previously isolated yeast genomic fragment complementing the *ole1* mutation has been sequenced and determined to encode the Δ9 fatty acid desaturase enzyme by comparison of primary amino acid sequence to the rat liver stearoyl-CoA desaturase. The *OLE1* structural gene encodes a protein of 510 amino acids (251 hydrophobic) having an approximate molecular mass of 57.4 kDa. A 257-amino acid internal region of the yeast open reading frame aligns with and shows 36% identity and 60% similarity to the rat liver stearoyl-CoA desaturase protein. This comparison disclosed three short regions of high consecutive amino acid identity (>70%) including one 11 of 12 perfect residue match. The predicted yeast enzyme contains at least four potential membrane-spanning regions and several shorter hydrophobic regions that align exactly with similar sequences in the rat liver protein. An *ole1* gene-disrupted yeast strain was transformed with a yeast-rat chimeric gene consisting of the promoter region and N-terminal 27 codons of *OLE1* fused to the rat desaturase coding sequence. Fusion gene transformants displayed near equivalent growth rates and modest lipid composition changes relative to wild type yeast control implying a significant conservation of Δ9 desaturase tertiary structure and efficient interaction between the rat desaturase and yeast cytochrome b_5.

In animal and fungal cells, monounsaturated fatty acids are synthesized via an aerobic process from saturated fatty acid precursors by a microsomal membrane-bound three-component enzyme system involving cytochrome b_5, NADH-dependent cytochrome b_5 reductase, and the Δ9 fatty acid desaturase (1-3). This complex catalyzes the insertion of a double bond between carbons 9 and 10 of the saturated fatty acyl substrates, palmitoyl (16:0)- and stearoyl (18:0)-CoA, yielding the monoenoic products palmitoleic (16:1) or oleic (18:1) acids. Although higher eukaryotes contain polyunsaturated fatty acids in their membranes, either synthesized endogenously via Δ12 and Δ15 desaturase reactions or obtained from their diet, the Δ9 reaction accounts for all *de novo* unsaturated fatty acid (UFA)[1] production in *Saccharomyces cerevisiae* (4).

Isolation and characterization of fatty acid desaturase enzymes has proved difficult due to their extraordinary hydrophobic nature and tight association with membranes. Although fatty acid desaturation was first described using the yeast Δ9 desaturase system, only animal Δ9 enzymes have been successfully purified to homogeneity (5, 6). At a genetic level, only the DNA sequence for the rat liver and mouse adipocyte genes have been reported and analyzed (7, 8). Those genes were found to encode proteins with 92% identical amino acid sequences.

The Δ9 desaturase from rat liver has been most extensively characterized. It is a protein consisting of 358 amino acids of which 62% are hydrophobic (7). The functional enzyme has an obligate phospholipid requirement and contains one molecule of non-heme iron (5). Effects of chemical modification on enzyme function has suggested that arginyl and tyrosyl residues are involved in the binding of the negatively charged CoA moiety of the substrate and in the chelation of the iron prosthetic group, respectively (9). A truncated rat liver Δ9 enzyme missing 26 residues from the N terminus is also membrane-bound and functional (10).

Yeast mutants bearing the *ole1* allele require oleic acid for growth and were believed to produce a defective Δ9 desaturase suggesting that *OLE1* was the structural gene encoding the enzyme (11). Recently, we isolated and characterized a yeast genomic fragment containing the *OLE1* gene of *S. cerevisiae* (12). Replacement of the wild type gene in haploid cells with a disrupted form of that fragment resulted in a UFA-requiring, nonreverting phenotype.

In this paper we report the DNA sequence of the *S. cerevisiae OLE1* gene and compare the deduced amino acid sequence of the yeast Δ9 fatty acid desaturase with that of the rat liver stearoyl-CoA desaturase primary sequence. Although the proteins encoded are highly divergent, the rat Δ9 desaturase gene functions efficiently in *S. cerevisiae* in place of the native yeast gene. Furthermore, predicted structural features of the two proteins suggest a model for the topology of the Δ9 fatty acid desaturase in the ER membrane.

* This work was supported by National Science Foundation Grant DMB84-17802, Biomedical Research Group Grant RR-07058-21 from the Public Health Service, and by a grant from the Bureau of Biological Research Charles and Johanna Busch Memorial Fund. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J05676.

‡ Supported by a Charles and Johanna Busch predoctoral fellowship.

§ Supported by an Arthur McCallum predoctoral fellowship.

¶ To whom correspondence should be addressed. Tel.: 201-932-4081 or 201-873-2752.

[1] The abbreviations used are: UFA, unsaturated fatty acid; ORF, open reading frame; ER, endoplasmic reticulum; kb, kilobase(s).

MATERIALS AND METHODS

DNA Manipulations, Media, and Stain—All recombinant DNA manipulations were according to standard methods (13, 43). Plasmid amplifications and bacterial transformations were performed using either *Escherichia coli* strain HB101 or XL1 Blue (Stratagene). Yeast transformations were by the method of Ito *et al.* (14). Growth analysis was performed in synthetic dextrose medium supplemented with the appropriate amino acids (23). The genotype of yeast strain L8-14C is: *a, ole1Δ::LEU2, leu2-3, leu2-d112, ura3-52, his4*).

DNA Sequencing—Overlapping DNA fragments lying within the *OLE1* open reading frame were subcloned into pBluescript vectors (Stratagene) in two orientations for sequencing in either direction. Single-stranded DNA squencing templates were prepared by methods supplied by Stratagene. The M13(−20) primer was hybridized to single-stranded DNA templates and DNA sequencing was performed by the dideoxy chain termination method of Sanger *et al.* (15) using the modified T7 DNA polymerase, Sequenase (U. S. Biochemical Corp.). In two cases, *OLE1* internal oligonucleotides were synthesized to facilitate DNA sequence analysis.

DNA Sequence and Deduced Primary Sequence Analysis—OLE1 DNA sequence and the deduced primary sequence analysis was performed using the Genetics Computer Group (GCG) sequence analysis software package (16). Amino acid sequence of the rat liver stearoyl-CoA desaturase was obtained from GenBank. Primary sequence comparison of the yeast and rat liver Δ9 desaturases was performed using the BestFit analysis program. Hydropathy analysis was according to Kyte and Doolittle (17).

Construction of Modified ole1 Alleles—Alleles *ole1-33* and *ole1-107*, containing stop codons in the 5′ region of the coding sequence, were constructed similarly. YEp352/OLE4.8 was partially digested with *Sal*I or *Nco*I, the cohesive ends were made blunt, and plasmids were religated. Following amplification in *E. coli*, plasmid samples were subject to restriction enzyme analysis. Candidates lacking the relevant restriction site were subject to DNA sequence analysis for verification.

Construction of Episomal and Centromeric Plasmids Bearing the Rat Liver Stearoyl-CoA Desaturase Gene—A 1.2-kb rat liver Δ9 desaturase cDNA fragment encoding residues 3–358, stop codon, and 136 base pairs of the 3′-untranslated region was removed from plasmid pDs3-358 (10) by digestion with *Bam*HI and *Sac*I and inserted into the multiple cloning site of episomal plasmid YEp352. A 1.0-kb yeast genomic fragment encompassing the promoter region, translation initiation codon, and the first 27 codons of the *OLE1* was isolated via *Hin*dIII/*Sal*I digestion and ligated in-frame with the rat desaturase fragment in YEp352. In this final construct, an eight-codon linker derived from the multiple cloning site regions of pUC8 and YEp352 separates the yeast N-terminal codons from the rat desaturase sequence. The predicted size of the fusion gene product is 391 amino acid residues. The yeast-rat fusion gene was then recovered via *Hin*dIII/*Dra*I digestion and ligated into YCp50 using *Hin*dIII and *Nru*I restriction sites. Plasmids bearing the fusion gene were amplified in *E. coli* strain XL1-Blue and used to transform the yeast *ole1* gene-disrupted strain L8-14C.

Lipid Isolation and Fatty Acid Analysis—Lipids were extracted from whole yeast cells by direct saponification (18). Fatty acid methyl esters were prepared by transmethylation with boron trifluoride (19) and analyzed by gas chromatography using a 30-meter capillary column SP-2330 (Supelco) in a Hewlett-Packard 5710A chromatograph as previously reported (12).

RESULTS AND DISCUSSION

General Features of the OLE1 Structural Gene—In a previous report it was shown that a cloned 4.8-kb *Hin*dIII yeast genomic fragment, but not two subclones of that fragment terminating at a central *Kpn*I region, complemented the *ole1* mutation of *S. cerevisiae* (12). From that *Kpn*I junction, overlapping subclones were used to "walk" through the observed open reading frame (ORF) in both directions yielding the sequence strategy presented in Fig. 1. Both strands were sequenced through the entire ORF without ambiguity.

The DNA and deduced amino acid sequence of *OLE1* and flanking nucleotide sequence is shown in Fig. 2. The ORF is 1530 nucleotides long. Translation of the entire ORF would produce a 510-amino acid polypeptide having an approximately molecular mass of 57.4 kDa containing 49.2% hydrophobic and 25.7% charged (10.0% acidic and 15.7% basic) amino acid residues. No consensus *N*-glycosylation sites are present in the deduced amino acid sequence of *OLE1* and the protein does not appear to contain a cleavable N-terminal signal sequence.

Yeast TATA promoter elements are commonly found 40–120 base pairs upstream from transcription initiation sites (20) with an average mRNA leader sequence of 52 nucleotides (21). The *OLE1* promoter region has two consensus TATA promoter elements (TATAAA and TATATA) located at positions −30 and −156 relative to the ORF. A transcription initiation event, directed from the TATATA element located at −156, could yield a transcript having features consistent with the above observations. However, transcription initiation directed from the TATAAA promoter element located at −30 could result in an atypically short, nontranslated leader sequence relative to the first in-frame ATG. Furthermore, there are three additional in-frame ATG codons within the first 400 base pairs of the *OLE1* ORF at positions 56, 61, and 116 that could also serve as potential translation start sites (see Fig. 1). Due to the close proximity of the first ATG codon to the TATA promoter element at −30 and comparison with the rat desaturase (discussed below) that showed no significant similarities in the first 140 amino acids, we were prompted to test for functional *OLE1* products initiating from these downstream sites. Two modified *ole1* alleles were constructed (see "Materials and Methods") that shifted the ORF and introduced translation stop codons at either position 33 (*ole1-33*) or 107 (*ole1-107*). Both in-frame stop codons were positioned before the next available ATG sequence. The *ole1* gene-disrupted yeast strain L8-14C, bearing the deletion allele *ole1Δ::LEU2*, was transformed with either of the above alleles on an episomal plasmid and tested for the ability to grow in the absence of exogenous UFAs. (Strains bearing this *ole1* allele were previously shown (12) to completely lack Δ9 desaturase activity as determined by product formation and have limited and finite growth potential (4–5 generations) in UFA-free medium.) In both cases the transformed strain grew only when UFAs were added to the growth medium, which is consistent with the first in-frame ATG codon functioning as the primary site of translation initiation.

Yeast and Rat Liver Δ9 Enzyme Amino Acid Analysis—A computer search of homologies to all current entries in GenBank/EMBL protein data bases identified a single data

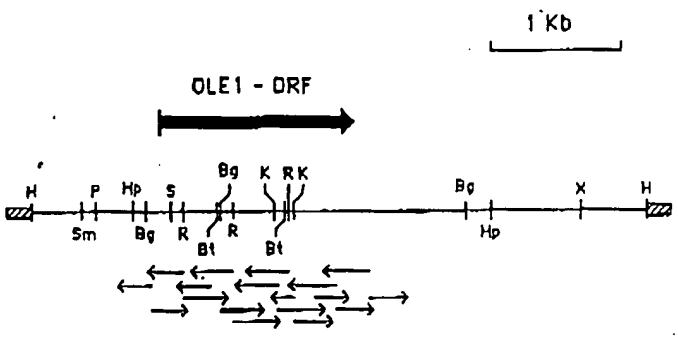


FIG. 1. ***OLE1* restriction map and sequencing strategy.** Primary restriction sites mapping the 4.8-kb yeast DNA fragment containing *OLE1*: *Bg*, *Bgl*II; *Bt*, *Bst*EII; *H*, *Hin*dIII; *Hp*, *Hpa*I; *K*, *Kpn*I; *R*, *Eco*RI; *S*, *Sal*I; *Sm*, *Sma*I; *P*, *Pst*I; and *X*, *Xho*I. The position and direction of the 1530-base-long ORF encoding the Δ9 enzyme is indicated by the *large arrow* above the map. *Small arrows* below the map indicate by size and direction the *OLE1* subclones used to sequence the entire ORF and flanking regions.

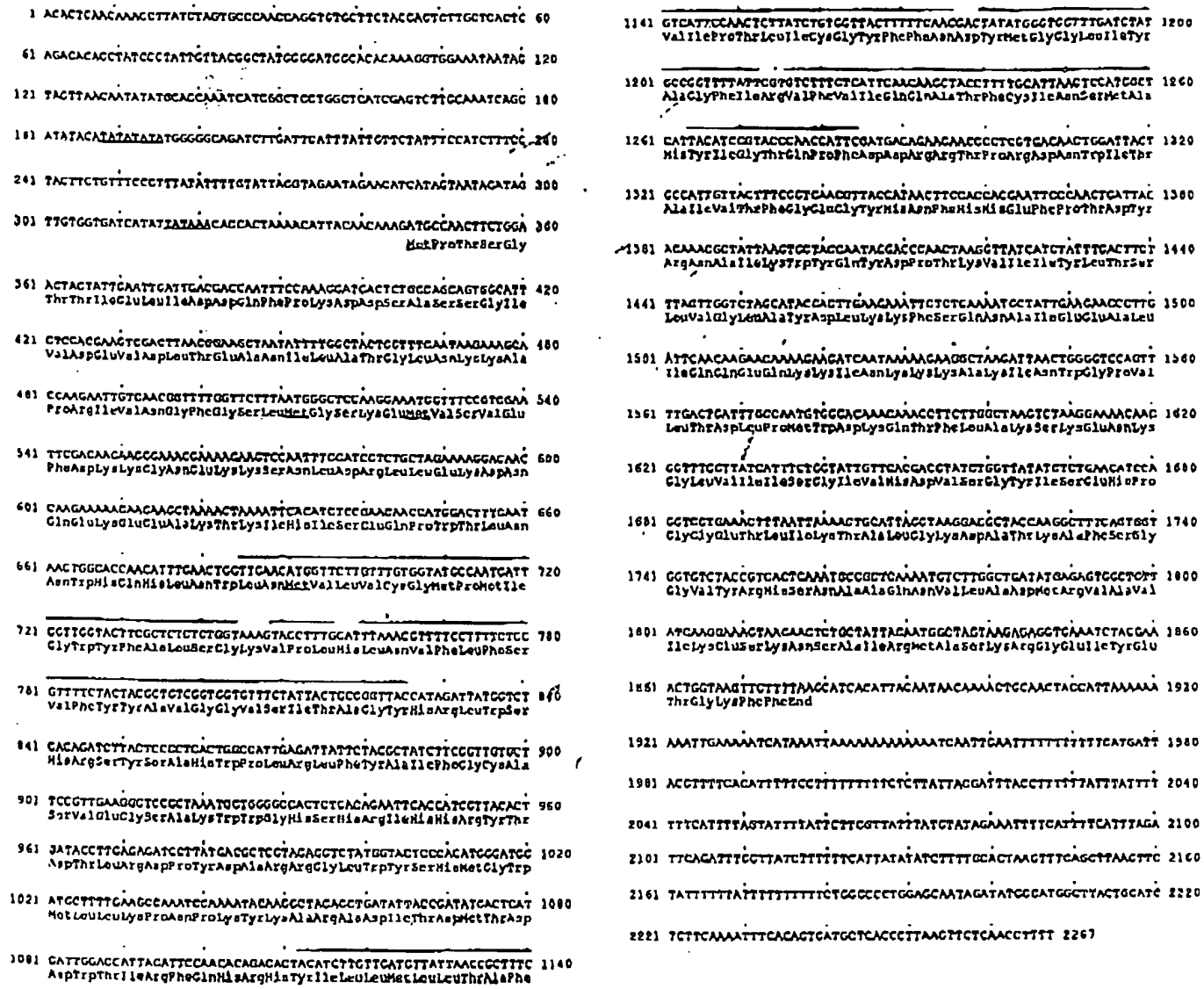


FIG. 2. Nucleotide and encoded amino acid sequence of the Δ9 fatty acid desaturase structural gene, *OLE1*. Two consensus yeast TATA elements preceding the 1530-base-long ORF and the first four in-frame methionine-specific codons are *underlined*. An *OLE1* internal region of 258 amino acids displaying significant identity to the rat liver Δ9 enzyme is delimited by *asterisks*. Potential membrane-spanning regions are highlighted with *lines* above nucleotide and amino acid sequences.

base entry, the rat liver Δ9 desaturase, with significantly similarity to the *OLE1* gene product (Fig. 3). The aligned sequences show 36% identity and are greater than 60% similar over the region encompassing the C-terminal 260 amino acids of the shorter rat protein. No significant similarities exist over the N-terminal 141 amino acids of the yeast and 99 amino acids of the rat sequences. The yeast open reading frame extends 113 amino acids beyond the C-terminal end of the rat sequence. Within the region of high amino acid similarity there are three segments, having a minimum length of 10 residues, of very high identity (>70%) beginning at *OLE1* amino acid positions 156, 331, and 368. The most highly conserved of these is the first region where 17 of 23 identities are observed including one stretch containing 11 of 12 perfect matches.

The most conserved amino acid type within the compared region of the yeast protein is histidine with 10 of 14 (71.4%) residues in perfect alignment. Two other amino acid residues, proline and arginine, also show greater than 50% total identity. Arginine residues of the rat liver enzyme have been

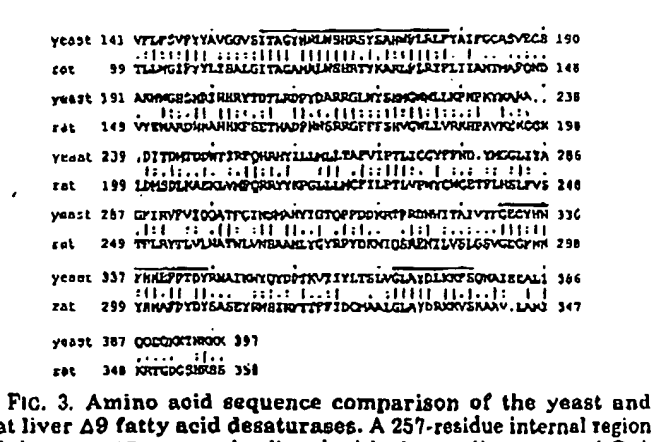


FIG. 3. Amino acid sequence comparison of the yeast and rat liver Δ9 fatty acid desaturases. A 257-residue internal region of the yeast Δ9 enzyme is aligned with the rat liver stearoyl-CoA desaturase. Comparison was prepared by the GCG sequence analysis program BestFit. Identical residue matches are indicated by *connecting solid lines*. Two or one point between residues indicate decreasing amino acid similarity. Percent similarity value is based on the number of identical and two-point amino acid comparisons. Segments showing

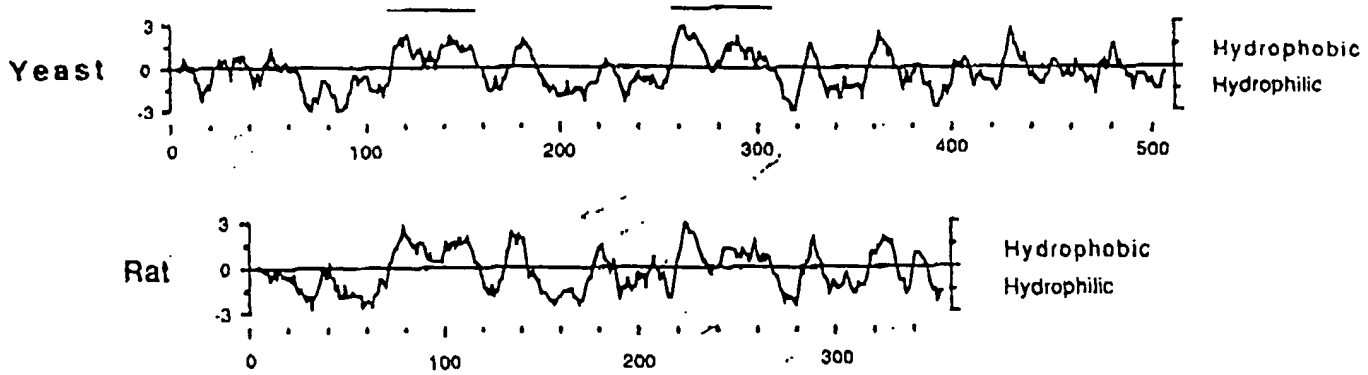


FIG. 4. **Hydropathy analysis of yeast and rat liver Δ9 enzymes.** Aligned Kyte-Doolittle (17) hydropathy profiles of the yeast and rat liver Δ9 desaturase proteins. The presumptive double membrane-spanning sequences are indicated by *bold lines* above those regions.

(9). Its significance as a highly conserved amino acid supports this finding. Although the role of histidyl residues in the fatty acid desaturases has not been examined, their highly conserved appearance also suggests an important contribution to enzyme function.

Structural Analysis and Proposed Topology of the Yeast Δ9 Enzyme—Striking similarities were also observed in the hydropathic characteristics of the two enzymes (Fig. 4). Both proteins contain two long hydrophobic regions (~50 residues) that could potentially form two membrane traversing loops, each consisting of two transmembrane segments. Chou-Fasman algorithms predict β-turn forming potential in the central portions of each loop in both the yeast and rat liver proteins. Inspection of the primary sequences at those sites reveals the presence of multiple helix-breaking amino acids that could serve to disrupt α-helical structure in order to form the looped structures. These hydrophobic regions are at identical positions in the aligned yeast and rat sequences. At least three smaller hydrophobic regions (each <7 amino acids) are also found at identical positions in the two proteins. The regions of high consecutive amino acid identity, however, are not within the hydrophobic sequences. The first region is located between the two "transmembrane loop regions," the second and third identity regions are located at the C-terminal part of the protein past the second "transmembrane loop." Neither extension of the N- and C-terminal domains of the yeast appears significantly hydrophobic and an examination of the amino acid distribution in those regions further suggests that they do not contribute to the integral membrane domains of the protein. A proposed model of the topology of the yeast protein in the ER membrane is given in Fig. 5. Assuming that the membrane-spanning regions are confined to the predicted hydrophobic sequences that are greater than 50 amino acids long, the arrangement places most of the protein on the cytosolic side of the ER membrane. Furthermore, all three regions of high consecutive identity would be located on that side of the membrane which is consistent with its proposed site of action (22).

Growth and Fatty Acid Content of Gene-disrupted Yeast Transformed with the Rat Liver Δ9 Desaturase—The significant sequence and predicted structural similarities observed between the yeast and rat Δ9 proteins prompted us to test whether the rat enzyme could functionally replace the yeast enzyme in *S. cerevisiae*, although there are additional residues at the N- and C-terminal ends of the yeast peptide sequence that are not found on the rat protein. A yeast-rat fusion gene was constructed (see "Materials and Methods") placing codons 3-368 of the rat gene in-frame with the initial 27 codons

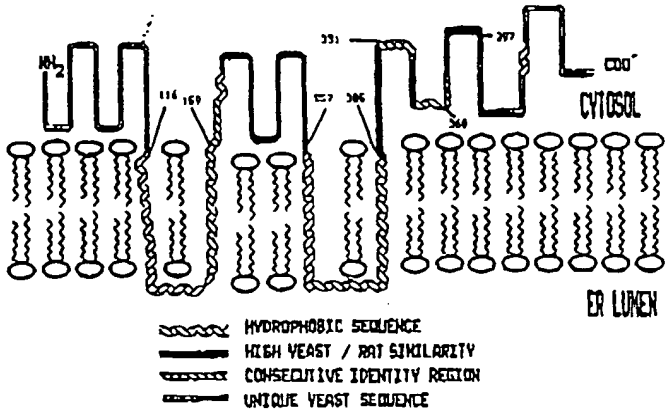


FIG. 5. **Model for the orientation of the yeast desaturase in the ER membrane.** The *numbers* identify amino acid positions in the yeast open reading frame.

of the yeast gene and promoter sequences separated by an 8-codon linker region. This fusion gene was placed on a multicopy episomal and single copy centromere-based (CEN) vectors and introduced into the *ole1* gene-disrupted yeast strain, L8-14C. Fusion gene transformants were analyzed for growth and lipid composition relative to the same gene-disrupted strain transformed with the plasmid bearing native *OLE1* gene.

Yeast transformants bearing either the native *OLE1* or the yeast-rat fusion desaturase gene (two isolates) on an episomal plasmid were found cured of the UFA requirement and, surprisingly, showed identical growth rates (Fig. 6*A*) indicating significant conservation of Δ9 desaturase tertiary structure and an ability of the rat enzyme to interact with the yeast cytochrome b_5. In addition, because the rat protein is 113 residues shorter than the yeast desaturase at the C-terminal end and yet can functionally substitute for the yeast enzyme in *S. cerevisiae*, it appears that this extension of the yeast protein may be nonessential for catalytic functions. We cannot exclude the possibility, however, that the additional residues may be involved in other functions that influence its catalytic efficiency or optimize interactions with other components of the desaturase system.

An analysis of stationary phase cellular lipid compositions revealed, however, significant differences in the percentage of 16-carbon fatty acid species in the yeast-rat fusion gene transformants relative to the wild type control and, as a result, a modest decrease in the percent total UFA (Table I). The lower percentage of 16:1 and increased 16:0 species found in

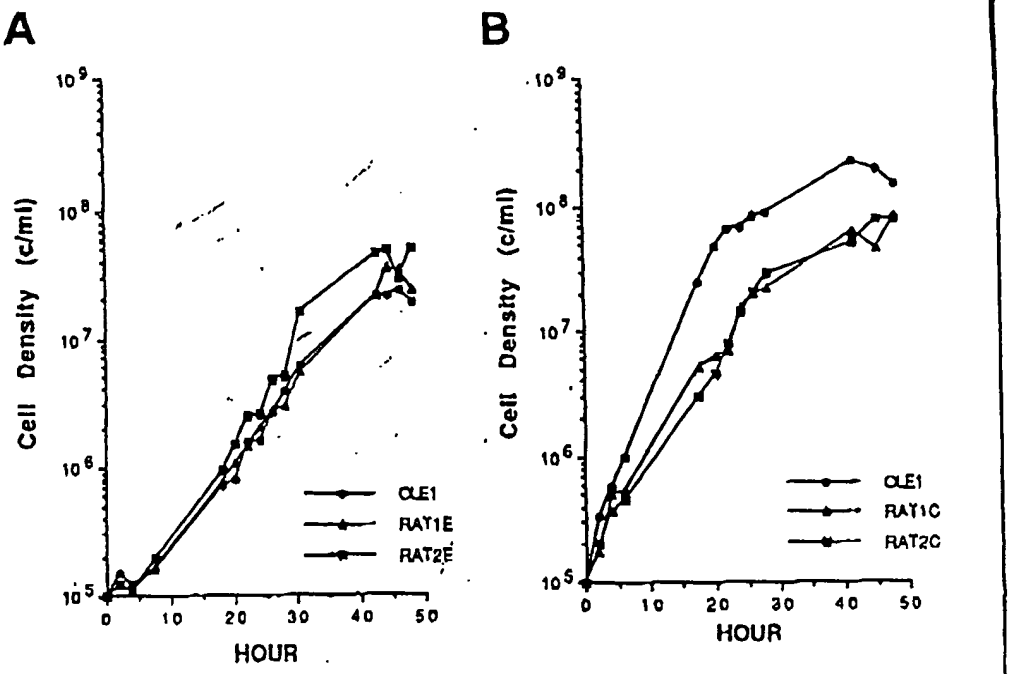


FIG. 6. Growth characteristics of transformed *S. cerevisiae*. Growth curves were determined for L8-14C transformants containing either the native *OLE1* gene or the yeast-rat Δ9 fusion gene on the episomal plasmid YEp352 (*A*) or the CEN plasmid YCp50 (*B*).

TABLE I

Fatty acid composition of transformed S. cerevisiae

Stationary phase L8-14C cells transformed with *OLE1* or the yeast-rat Δ9 chimeric gene on multiple (episomal) or single (CEN) copy number plasmids were harvested and cellular lipids analyzed as described under "Materials and Methods."

Plasmid type and transformant	Fatty acids					
	14:0	16:0	16:1	18:0	18:1	UFA
	%					
Episomal						
OLE1	1.95	21.92	41.70	4.65	29.77	71.47
RAT1E	1.40	33.00	28.77	4.61	32.21	60.98
RAT2E	1.41	30.94	27.42	4.73	35.51	62.93
CEN						
OLE1	1.16	18.38	34.05	9.10	37.32	71.36
RAT1C	5.58	40.67	21.22	8.70	23.82	45.04
RAT2C	7.23	39.71	15.83	10.89	26.35	42.18

those strains may reflect a preference of the rat Δ9 enzyme for the 18:0-CoA substrate over 16:0-CoA.

Although a yeast-rat Δ9 desaturase fusion gene is capable of functionally replacing the native *OLE1* of *S. cerevisiae* when present on a high copy number plasmid, a more stringent test of the efficiency of the rat protein in yeast would be to examine cells transformed with a single copy of the fusion gene. Cells containing the chimeric gene on CEN plasmid YCp50 showed growth rates that are reduced approximately 65% relative to wild type (Fig. 6*B*).

Similarly, the lipid composition of CEN plasmid-bearing yeast transformants differed markedly between those containing the chimeric gene and those containing the cloned yeast gene (Table I). The relative UFA levels were reduced approximately 38% in cells containing the rat gene coding sequence and the compensatory relative increase in saturated fatty acids resulted in a doubling of the 16:0 content and increased 14:0 levels, but no significant change in the level of 18:0. Thus, the yeast-rat Δ9 desaturase fusion gene can functionally replace the native *OLE1* of *S. cerevisiae*, although its action results in striking differences in cellular fatty acid compositions.

In previous studies using gene disruption and lipid analytical methods (12) we provided evidence suggesting that the *OLE1* gene encoded the yeast Δ9 fatty acid desaturase. The deduced *OLE1* amino acid sequence and physical comparisons of the yeast and rat liver proteins given here provide further proof that the *OLE1* locus contains the authentic structural gene for the desaturase. The aligned regions of consecutive identity between these two proteins from widely divergent sources suggests that they may represent conserved regions with similar function. The finding that the rat Δ9 fatty acid desaturase gene can complement *OLE1* in *S. cerevisiae* although the two proteins have only 36% identity suggests that there is conserved functional interaction among cytochrome b_5-mediated desaturase systems. Thus, ER-bound Δ9 enzymes from other organisms and possibly other cytochrome b_5-mediated desaturases, such as the Δ12 and Δ15, may also function in yeast.

Acknowledgment—We wish to thank Philipp Strittmatter for plasmids containing the rat stearoyl-CoA desaturase gene.

REFERENCES

1. Bloomfield, D. K., and Bloch, K. (1960) *J. Biol. Chem.* 235, 337-345
2. Tamura, Y., Yoshida, Y., Sato, R., and Kumaoka, H. (1976) *Arch. Biochem. Biophys.* 175, 284-294
3. Ohba, M., Sato, R., Yoshida, Y., Bieglmayer, C., and Ruis, H. (1979) *Biochim. Biophys. Acta* 572, 352-362
4. Schweizer, E. (1984) in *Fatty Acid Metabolism and Its Regulation* (Numa, S., ed) pp. 59-83, Elsevier Scientific Publishing Co., New York
5. Strittmatter, P., Spatz, L., Corcoran, D., Rogers, M. J., Setlow, B., and Redline, R. (1974) *Proc. Natl. Acad. Sci. U. S. A.* 71, 4565-4569
6. Prasad, M. R., and Joshi, V. C. (1979) *J. Biol. Chem.* 254, 6362-6369
7. Thiede, M. A., Ozols, J., and Strittmatter, P. (1986) *J. Biol. Chem.* 261, 13230-13235
8. Ntambi, J. M., Buhrow, S. A., Kaestner, K. H., Christy, R. J., Sibley, E., Kelly, T. J., Jr., and Lane, M. D. (1988) *J. Biol. Chem.* 263, 17291-17300
9. Enoch, H. G., and Strittmatter, P. (1978) *Biochemistry* 17, 4927-

4932
10. Strittmatter, P., Thiede, M. A., Hackett, C. S., and Ozols, J. (1988) *J. Biol. Chem.* **263**, 2532–2535
11. Resnick, M. A., and Mortimer, R. K. (1966) *J. Bacteriol.* **92**, 597–600
12. Stukey, J. E., McDonough, V. M., and Martin, C. E. (1989) *J. Biol. Chem.* **264**, 16537–16544
13. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
14. Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) *J. Bacteriol.* **153**, 163–168
15. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
16. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395
17. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132
18. Kates, M. (1972) *Techniques of Lipidology*, Elsevier Scientific Publishing Co., New York
19. Morrison, W. R., and Smith, L. M. (1964) *J. Lipid Res.* **5**, 600–608
20. Struhl, K. (1989) *Annu. Rev. Biochem.* **58**, 1051–1077
21. Cigan, A. M., and Donahue, T. F. (1987) *Gene (Amst.)* **59**, 1–18
22. Jeffcoat, R., Brawn, P. R., Safford, R., and James, A. T. (1977) *Biochem. J.* **161**, 431–437
23. Sherman, F., Fink, G. R., and Hicks, J. B. (1982) *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

In re: Meyers
Appl. No. 09/464,039
Filed December 15, 1999

APPENDIX D

Biochem. J. (1998) 332, 153–159 (Printed in Great Britain)

Saccharomyces cerevisiae GPI10, the functional homologue of human *PIG-B*, is required for glycosylphosphatidylinositol-anchor synthesis

Christine SÜTTERLIN*, M. Victoria ESCRIBANO†, Peter GEROLD‡, Yusuke MAEDA§, Maria J. MAZON†, Taroh KINOSHITA§, Ralph T. SCHWARZ‡ and Howard RIEZMAN*[1]

*Biozentrum of the University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland, †Instituto de Investigaciones Biomedicas, 28029 Madrid, Spain, ‡Philipps-Universität Marburg, Robert-Koch-Str. 17, 35037 Marburg, Germany, and §Department of Immunoregulation, Research Institute of Microbial Diseases, Osaka University, Osaka 565, Japan

An increasing number of plasma membrane proteins have been shown to be attached to the membrane via a glycosylphosphatidylinositol (GPI) moiety. All eukaryotes share a highly conserved GPI-core structure EthN-P-Man_3-GlcN-PI, where EthN is ethanolamine. We have identified a protein encoded by the yeast open reading frame YGL142C that shares 33% identity with the human Pig-B protein. Deletion of this essential gene leads to a block in GPI anchor biosynthesis. We therefore named the gene *GPI10*. Gpi10p and Pig-B are functional homologues and the lethal deletion of *GPI10* can be rescued by expression of the *PIG-B* cDNA. As found for *PIG-B* mutant cells, *gpi10* deletant cells cannot attach the third mannose in an α-1,2 linkage to the GPI core-structure intermediate. Overexpression of *GPI10* gives partial resistance to the GPI-synthesis inhibitor YW3548, suggesting that this gene product may affect the target of the inhibitor.

INTRODUCTION

Glycosylphosphatidylinositol (GPI) anchoring represents a mechanism for attachment of proteins to membranes found in all eukaryotic cells [1,2]. All eukaryotes share the common GPI core structure EthN-P-6Manα-1,2Manα-1,6Manα1,4-GlcNα-1,6-*myo*-inositol-P-lipid, where EthN is ethanolamine. Modifications of this core structure can occur on different mannose residues and depends on the species as well as on its developmental stage. The lipid moiety of the GPI anchor is also subject to modifications, such as variations of the fatty acids, ceramide remodelling or the attachment of alkyl chains [1].

The complete GPI precursor, whose structure has been solved for some species, is assumed to be assembled by sequential addition of the sugar components and ethanolamine to phosphatidylinositol [3]. The complete precursor is transferred *en bloc* to the anchor-attachment site of the protein (ω-site) with concomitant release of a C-terminal peptide [4]. Anchor attachment is thought to occur as a transamidation reaction [5,6]. A protein to be GPI-anchored carries two signal sequences: a cleavable N-terminal signal sequence that causes the protein to be translocated into the lumen of the endoplasmic reticulum (ER), and a C-terminal signal sequence that directs attachment of the GPI anchor. The consensus sequence for GPI-anchor addition consists of an amino acid with a small side-chain at the ω-site, two amino acids with small side-chains in positions ω+1 and ω+2, and a short hydrophilic spacer region followed by a hydrophobic domain of about 15–20 amino acids [7].

The identification of at least eight complementation classes of mutants from murine lymphoma and human K562 cells, lacking the expression of GPI-anchored proteins at the cell surface, led to the characterization of genes involved in mammalian anchor biosynthesis [3]. Three complementation classes, A, C and H, are blocked in the first step of precursor synthesis, which is the addition of GlcNAc to PI [8]. Class J mutants and class L mutants are defective in deacetylating GlcNAc-PI to give GlcN-PI [9–11]. Class E mutants are defective in dolichyl (Dol)-P-Man synthase, thereby affecting GPI mannosylation and N-linked glycosylation [12]. Class B mutant cells are blocked in the addition of the third mannose to the intermediate structure [13], the class F mutation affects the transfer of the terminal ethanolamine to the third mannose [12], and class K mutants accumulate the complete precursor without transferring it to the protein [10]. Whereas in mammalian cells, several cDNAs that encode components of the GPI synthesis machinery (*PIG-A*, *PIG-H*, *PIG-L*, *PIG-F*, *PIG-C* and *PIG-B*) have been cloned [11,13–17], only genes involved in the first step of anchor synthesis (*GPI1*, *GPI2*, *GPI3*/*SPT13*) and in anchor attachment (*GAA1*, *GPI8*) have been identified from yeast [18–21]. Yeast mutants that affect various steps of GPI-anchor synthesis have also been described [22].

The core structure of the GPI anchor contains three mannose residues which are all assumed to be transferred from Dol-P-Man as donor [23]. The mannose residues were found to be attached in different linkages, suggesting that at least three gene products are involved in GPI mannosylation. The Pig-B protein is required for the transfer of the third mannose in an α-1,2 linkage to the core structure, but to date, no enzymic activity of the Pig-B protein has been shown [13]. A yeast mutant, blocking a similar step as the Pig-B mutant, has been isolated [22]. Addition of the third mannose can also be blocked using a species-specific inhibitor, YW3548 [24]. YW3548 blocks GPI-anchor synthesis in mammalian cells and in yeast, but not in protozoa.

In this study, we report the identification of a yeast gene, *GPI10*, which appears to be the functional homologue of the human *PIG-B* gene. In order to study the function of this essential gene, we placed *GPI10* under the control of the GAL1/10 promoter. When *GPI10* expression was turned off, we found that inositol was no longer incorporated into proteins and that the maturation of Gas1p, a major GPI-anchored protein from yeast, was strongly reduced. Under these conditions, we observed the accumulation of the GPI intermediate Man_4-GlcN-(acyl)PI, and of a novel yeast GPI-derived structure that was sensitive to treatment with HF and to Jack bean mannosidase,

Abbreviations used: ER, endoplasmic reticulum; EthN, ethanolamine; GPI, glycosylphosphatidylinositol; GU, glucose unit.

[1] To whom correspondence should be addressed.

most likely Man-(EthN-P)Man-GlcN-(acyl)PI. The appearance of this novel lipid was blocked by addition of YW3548, a GPI-synthesis inhibitor. Overexpression of *GPI10* rendered wild-type cells partially resistant to YW3548.

MATERIALS AND METHODS

Strains and growth conditions

Cells were grown to saturation in either YPG [3% (v/v) glycerol, 2% (w/v) peptone, 1% (w/v) yeast extract, 40 mg/l each of adenine, uracil and tryptophan] or in SGYE [3% (v/v) glycerol, 0.67% (w/v) yeast nitrogen base, 0.2% (w/v) yeast extract and the required nutrients] and used to inoculate the glucose-containing media SDYE [0.2% (w/v) yeast extract, 0.67% (w/v) yeast nitrogen base and 5% (w/v) glucose] or YPUAD [1.0% (w/v) yeast extract, 2% (w/v) peptone, 5% (w/v) glucose and 40 mg/l each of uracil and adenine]. For overexpression experiments, cells were grown in YGal medium [2% (w/v) galactose, 1% (w/v) yeast extract, 2% (w/v) peptone, 40 mg/l of each of adenine, uracil and tryptophan]. Sporulation of yeast strains was on minimal sporulation medium (1% potassium acetate, 2% Bacto-agar, and the required nutrients to complement auxotrophies).

The full-length sequence of *GPI10* was replaced by the KanMX module using a PCR-based strategy [25]. A knockout cassette was constructed using a short flanking homology strategy and the oligomers: (upstream) 5′ TCTCTCAACGATAGGGTCT-GATTTATTAATTTTACTACTGCCGAAACGTACGCTGC-AGGTCGACC 3′ and (downstream) 5′ ATCGATGAATT-CGAGCTCGTTATGGTAAGTTAATATCGCTATAAGGT-CGCCGTCTATAAATTTAAA 3′.

This cassette was transformed into the diploid strains FY1679 and CEN.PK2. Deletion of the gene was monitored by resistance to geneticin (Gibco-BRL) and resistant colonies were confirmed by PCR analysis. For overexpression studies from the GAL1 promoter, RH1657 was used. For galactose-depletion studies, RH3699 was transformed with pGal-GPI10 and tetrads were dissected, generating spore A (RH3994), spore B (RH3995), spore C (RH3996) and spore D (RH3997). For *in vivo* mannose labelling experiments, RH3998 and RH3999 were used. All yeast strains used in this study are shown in Table 1.

Plasmids

pGal-GPI10 was constructed by PCR amplification of the full-length *GPI10* coding sequence from genomic DNA, with the primers (upstream) 5′ GATAGTCTAGAATTGTGTAATC CAA 3′ and (downstream) 5′ GTAAACTGATAAGTGAAG-CATGCAC 3′, using PFU Polymerase (Stratagene). The PCR fragment containing sites for *Xba*I (upstream oligomer) and *Sph*I (downstream oligomer) was subcloned into the *Xba*I–*Sph*I site of the pSEYC68 vector containing the GAL1/10 promoter cloned into the *Eco*RI–*Bam*HI site (referred to as pGal).

p425-PIG-B was prepared by subcloning the full-length *PIG-B* cDNA [13] into the *Eco*RV site of pBluescriptII (Stratagene), and the fragment generated by cutting with *Spe*I and *Sal*I was ligated into p425 digested with *Spe*I and *Sal*I.

[^{3}H]*myo*-Inositol labelling of proteins

Wild-type cells or *gpi10*Δ::KanMX cells carrying pGal-GPI10 were grown in SGYE to saturation and used to inoculate SDYE (5% glucose) in which they were grown for about 16 h. Cells (5×10^7) were washed twice in SD-inositol medium (5% glucose) [26], resuspended in 500 μl of SD without inositol (5% glucose) and depleted of inositol for 10 min before the addition of 15 μCi of [^{3}H]*myo*-inositol (Dupont de Nemours, Bad Homburg, Germany). Cells were labelled for 30 min. A total protein extract was prepared by lysing the cells in TEPI [100 mM Tris/HCl, pH 7.5/10 mM EDTA/proteinase inhibitors (1 μg/ml pepstatin, 1 μg/ml leupeptin, 1 μg/ml antipain)] by vortexing for 4 × 1 min with glass beads. Total protein was precipitated with trichloroacetic acid and the precipitates were washed with acetone and resuspended in Laemmli protein sample buffer. The proteins were separated by SDS/PAGE (10% acrylamide) and, after incubation of the gel in 1 M sodium salicylate, the radioactivity was detected by fluorography.

Pulse–chase analysis of Gas1p and CPY

Pulse–chase labelling and analysis of immunoprecipitates was done as described previously [26]. Briefly, wild-type cells and *gpi10*Δ::KanMX cells carrying pGal-GPI10 were grown to saturation in SGYE and shifted to SDYE (5% glucose) for 16 h. Cells were harvested and washed with SD* [5% (w/v) glucose, 0.67% (w/v) yeast nitrogen base and the required nutrients]. Cells (2.5×10^7 per time point) were resuspended in 1 ml of SD*, preincubated for 10 min, labelled with TRANS ^{35}S-label (Dupont de Nemours) and chased for the indicated times with 0.003% (w/v) methionine/0.003% (w/v) cysteine/3 mM $(NH_4)_2SO_4$. Aliquots were taken and the chase was stopped by adding NaN_3 and NaF to a final concentration of 10 mM. Cells were lysed by vortexing for 4 × 1 min with glass beads in TEPI. The lysates were boiled in the presence of 1% SDS for 5 min and centrifuged for 15 min in an Eppendorf centrifuge. The supernatant was added to 5 ml of TNET (100 mM Tris/HCl, pH 8/100 mM NaCl/5 mM EDTA/1% Triton X-100) and the extracts were incubated with polyclonal antisera against Gas1p or CPY and

Table 1 Strains of *Saccharomyces cerevisiae* that were used for this study

Strain	Phenotype	Source
FY1679	Mata/α *ura3-52/ura3-52 leu2Δ1/+ trp1Δ63/+* his3Δ200/+	Eurofan
Cen.PK2	Mata/α *ura3-52/ura3-52 leu2-3,112/leu2-3,112 trp1-289/trp1-289 his3Δ1/his3Δ1*	Eurofan
RH3698	FY1679, but *gpi10*Δ::KanMX/*GPI10*	This study
RH3699	Cen.PK2, but *gpi10*Δ::KanMX/*GPI10*	This study
RH3994	Mata *ura3-52 leu2-3,112 trp1-289 his3*Δ pGal-GPI10	This study
RH3995	Matα *gpi10*Δ::KanMX *ura3-52 leu2-3,112 trp1-289 his3*Δ pGal-GPI10	This study
RH3996	Mata *ura3-52 leu2-3,112 trp1-289 his3*Δ pGal-GPI10	This study
RH3997	Matα *gpi10*Δ::KanMX *ura3-52 leu2-3,112 trp1-289 his3*Δ pGal-GPI10	This study
RH3998	Matα *pmi40 ura3 leu2 ade2 trp1 his4* pGal-GPI10	This study
RH3999	Matα *pmi40 gpi10*Δ::KanMX *ura3 leu2 ade2 trp1 his4* pGal-GPI10	This study
RH1657	Mata *his4 leu2 ura3 lys2 bar1*	Laboratory strain

Protein A–Sepharose (Pharmacia, Uppsala, Sweden) for 3 h at room temperature. The immunoprecipitates were washed four times with TNET, once with 20 mM Tris, pH 7.4, resuspended in Laemmli sample buffer and analysed by SDS/PAGE, followed by exposure and quantification of the gel on a PhosphorImager (Molecular Dynamics, Sunnyvale, CA, U.S.A.). The percentage of mature Gas1p was determined as the percentage of the total Gas1p signal found in the 125 kDa form.

Glycolipid labelling *in vivo*

[^{14}C]Mannose labelling experiments *in vivo* using RH3998 and RH3999 were performed as described [27]. In brief, *pmi40* or *pmi40 gpi10*Δ::KanMX cells carrying the plasmid pGal-GPI10 were grown for 16 h in SDCU medium [5% (w/v) glucose/1% (w/v) peptone/ 0.67% (w/v) yeast nitrogen base/0.1% (w/v) mannose, supplemented with 40 mg/l uracil]. Cells (3×10^7) were resuspended in SPCU medium [0.1% (w/v) glucose, 2% (w/v) pyruvate, 0.67% (w/v) yeast nitrogen base and the required nutrients], preincubated at a non-permissive temperature for *pmi40* for 20 min and labelled with 25 μCi of [^{14}C]mannose for 45 min. In some experiments, cells were incubated with 5 μg/ml YW3548 (Novartis AG, Basel, Switzerland) for 10 min before addition of the radioactivity. The reaction was stopped by the addition of 10 mM NaF/10 mM NaN_3, and lipids were extracted with $CHCl_3/CH_3OH/H_2O$ (10:10:3, by vol.). Lipids were desalted by phase partitioning between n-butanol and 0.1 mM EDTA/5 mM Tris/HCl, pH 7.5, and analysed by TLC using $CHCl_3/CH_3OH/H_2O$ (10:10:3, by vol.) as solvent. For structural analysis, the accumulated lipids were extracted from the TLC plate, converted into hydrophilic fragments or core glycans and analysed by Bio-Gel P4 or high-pH anion-exchange chromatography respectively, as described previously [28,29]. The hydrophilic fragment generated using HNO_2 was digested with Jack bean mannosidase (Sigma) for 40 h with 2×1.5 units. The resulting fragment was desalted and purified on a G-15 column, then analysed on Dionex as above.

Halo assay

RH1657 cells expressing either pGal or pGal-GPI10 were grown to saturation in YGal medium. Cells (5×10^6) were included in 10 ml agar plates (0.8% agar) of the same media and 4 μl of YW3548 at different concentrations was spotted on the plate.

RESULTS

Identification of *GPI10*

A database search using the National Center of Biotechnology Information BLAST server indicated a large extent of homology between the protein encoded by the yeast open reading frame YL142C and the human Pig-B protein. Owing to the results presented below, we named this gene *GPI10*. The *GPI10* gene encodes a protein of 616 amino acids with a predicted size of about 72 kDa. It shows 33% identity with and 58% similarity to the human *PIG-B* gene (Figure 1), the potential third mannosyltransferase producing the GPI anchor. In addition, the hydrophobicity profiles of the two proteins are very similar (results not shown).

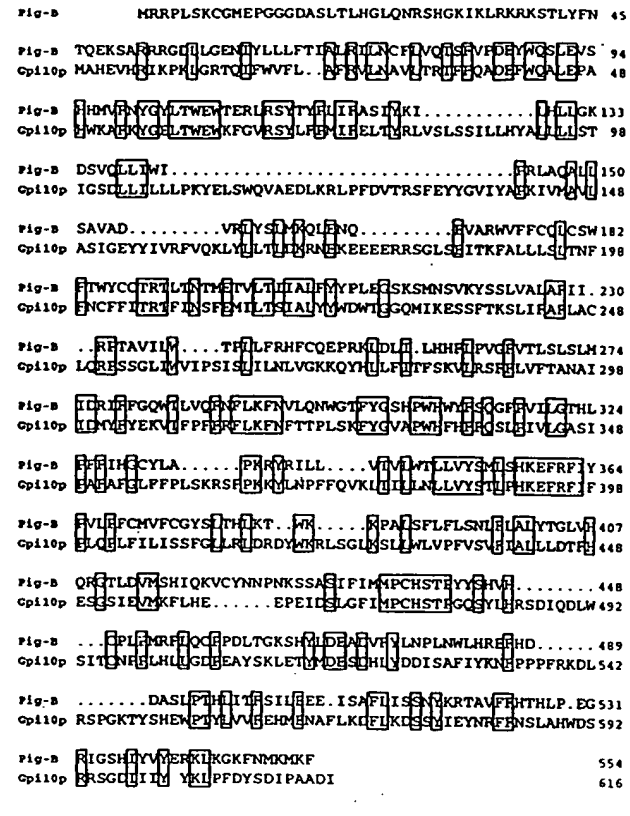


Figure 1 Sequence alignment of Gpi10p and the Pig-B protein

The sequence alignment was done using the BESTFIT program of the Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI, U.S.A. Identical residues are shown in boxes.

GPI10 is essential and its deletion can be rescued by expression of the human *PIG-B* cDNA

To test whether *GPI10* is an essential gene, we constructed a KanMX knockout cassette and transformed it into the diploid wild-type strain Cen.PK2 to disrupt the *GPI10* gene [25]. Upon tetrad dissection, we found that only two spores from each tetrad could grow (Figure 2A). The growing spores were sensitive to geneticin, showing that spores that were deleted for *GPI10* were inviable. *GPI10* was also found to be essential in other strain backgrounds (results not shown).

To study the function of the *GPI10* gene product, we constructed a conditional mutant of *GPI10* by cloning the gene under control of the GAL1/10 promoter on a centromeric plasmid. In this way, the expression of *GPI10* can be turned off simply by shifting the cells from glycerol-containing medium (SGYE or YPG), in which the promoter is not repressed, to glucose-containing medium (SDYE or YPD, both containing 5% glucose) where it is. The Gal construct of *GPI10* (pGal-GPI10) was functional and could restore growth of *gpi10*Δ::KanMX spores (Figure 2B). This construct was used to perform Gpi10p depletion studies.

GPI10 had been identified based on its high degree of homology to the human *PIG-B* gene. We therefore tested whether *GPI10* could be functionally replaced by *PIG-B*. *PIG-B* was cloned into a yeast multicopy vector (p425) behind the MET promoter, which is activated in the absence of methionine. p425-PIG-B was transformed into the Cen.PK2 diploid strain which is heterozygous for *gpi10*Δ::KanMX. The diploid was sporulated on minimal sporulation plates and tetrads were dissected onto minimal medium plates without methionine. Despite low overall spore viability, we obtained spores that carried the plasmid and

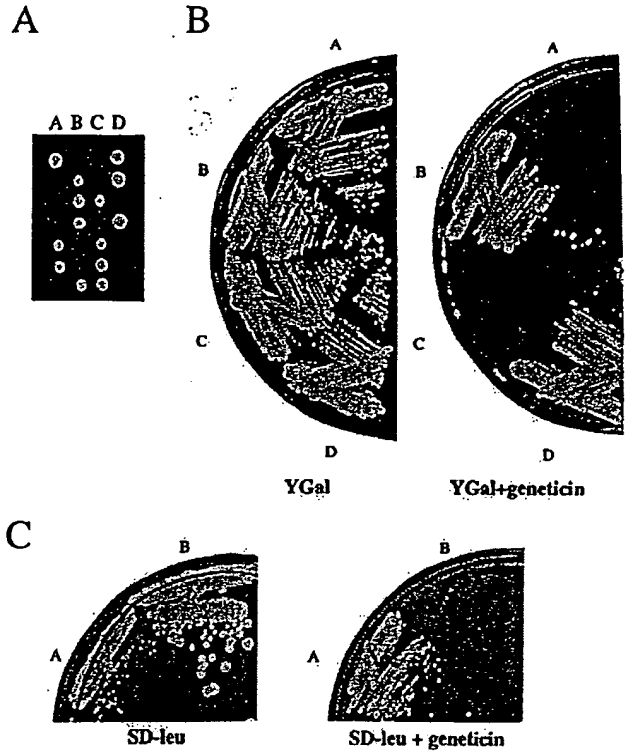


Figure 2 *GPI10* is essential for growth and its deletion can be complemented by pGal-GPI10 or by p425-PIG-B

(A) The *GPI10* open reading frame was replaced by KanMX in the diploid Cen.PK2 and tetrads were dissected. (B) pGal-GPI10 was transformed into RH3699 and tetrads were dissected. Growth on medium containing geneticin was used to monitor the deletion of *GPI10*. (C) p425-PIG-B was transformed into RH3699 and tetrads were dissected onto minimal medium. The growth of two spores (A, B, C and D) is shown: spore A was disrupted for *GPI10*, as monitored by growth on medium containing geneticin, whereas the wild-type spore, B, cannot grow on geneticin-containing plates.

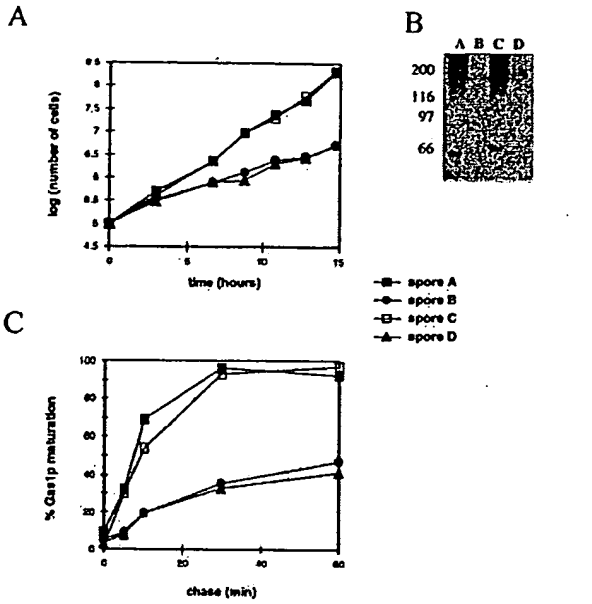


Figure 3 Gpi10p depletion causes a GPI-anchoring defect

(A) The growth of the spores of one tetrad [RH3994 (A), RH3995 (B), RH3996 (C), RH3997 (D)] was monitored upon shift from glycerol- to glucose-containing medium. (B) Incorporation of [^{3}H]*myo*-inositol into proteins is shown for the spores of the tetrad. Cells were shifted from glycerol- to glucose-containing medium for 16 h. They were then labelled with [^{3}H]*myo*-inositol, followed by preparation of a total cell lysate, analysis by SDS/PAGE and fluorography. (C) Pulse–chase experiments were performed in the spores of the tetrad. Cells were shifted from glycerol- to glucose-containing medium for 16 h. They were then metabolically labelled and chased as indicated. Lysates were prepared which were immunoprecipitated for Gas1p. The signal was detected by exposure and quantification of SDS gels using a PhosphorImager. Quantification of these pulse–chase experiments are shown. The appearance of mature Gas1p indicates GPI-anchor attachment and transport to the Golgi apparatus.

which were resistant to geneticin (Figure 2C), suggesting that they were deleted for *GPI10*. *PIG-B* expression did not completely restore wild-type growth levels in *gpi10*Δ::KanMX strains [compare spore A (*gpi10*Δ::KanMX) and spore B (wild type) in Figure 2C], but partial restoration of phenotypes is commonly seen for functional complementation over such a large evolutionary distance. Expression of *PIG-B* from a centromeric vector was not sufficient to restore growth (results not shown). We conclude that expression of the human *PIG-B* cDNA can complement the lethal phenotype of *gpi10*Δ::KanMX cells, suggesting that the two proteins are functional homologues.

Gpi10p depletion results in a GPI-synthesis defect

The GAL1/10-controlled *GPI10* vector allowed us to regulate the expression of *GPI10*. To find conditions where Gpi10p was functionally depleted, we followed the growth of the spores of a tetrad (Figure 2B) upon shift from glycerol- to glucose-containing medium. As shown in Figure 3(A), the growth of *gpi10*Δ::KanMX cells slowed down significantly after about 8 h incubation in glucose-containing medium, suggesting that Gpi10p was strongly depleted. For the following experiments, we used conditions where *GPI10* expression was turned off for 8–16 h. We had also tried to shift cells from galactose- to glucose-containing medium, but observed no change in growth rate over a period of 24 h (results not shown). This could be due to a very high expression of Gpi1p on galactose medium combined with a long half-life of the protein.

Since *GPI10* appears to be the functional homologue of *PIG-B*, we tested whether *GPI10* is indeed involved in GPI-anchor synthesis. GPI-anchored proteins are the only proteins that incorporate [^{3}H]*myo*-inositol due to the inositol residue in the GPI anchor. Cells were grown in SGYE to saturation and used to inoculate SDYE (5% glucose). After 16 h growth in SDYE (5% glucose) the cells were labelled for 30 min with [^{3}H]*myo*-inositol, total protein was extracted and analysed by SDS/PAGE, followed by fluorography. We observed that in *gpi10*Δ::KanMX spores (B and D), the incorporation of radiolabelled inositol was greatly reduced when compared with the wild-type spores (A and C), demonstrating a defect in GPI-anchoring (Figure 3B). Furthermore, pulse–chase analysis of the GPI-anchored protein Gas1p was performed. Gas1p occurs as a 105 kDa form in the ER. Upon arrival at the Golgi, its core glycans are elongated, causing a shift in apparent molecular mass from 105 to 125 kDa. Transport of Gas1p from the ER to the Golgi depends on addition of the GPI anchor [30,31]. After 16 h of shift from SGYE medium to SDYE medium (5% glucose), cells were labelled for 5 min and chased for various amounts of time. Cells were lysed and either Gas1p or the non-GPI-anchored vacuolar hydrolase carboxypeptidase Y was precipitated from total lysates using polyclonal antisera. Immune complexes were separated by SDS/PAGE and analysed on a PhosphorImager. As shown in

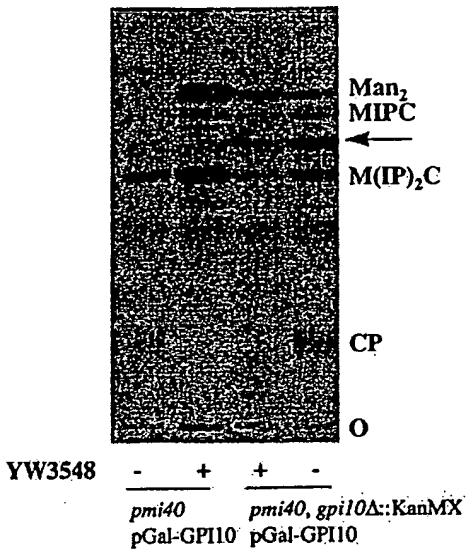


Figure 4 Gpi10p depletion causes the accumulation of Man_2-GlcN-(acyl)PI and of a novel lipid

pmi40 (RH3998) or *pmi40 gpi10*Δ::KanMX (RH3999), both carrying pGal-GPI10, were shifted to glucose-containing medium for 16 h. Cells were labelled for 45 min with [^{14}C]mannose, followed by extraction of total lipids and analysis by TLC. Man_2, Man_2-GlcN-(acyl)PI; MIPC, mannosylinositolphosphorylceramide; $M(IP)_2C$, mannosyldi-inositolphosphorylceramide; arrow, novel lipid; CP, complete precursor; O, origin.

Figure 3(C), Gas1p maturation was severely delayed in *gpi10*Δ::KanMX spores (B and D) when compared with wild-type spores (A and C). Under these conditions, maturation of carboxypeptidase Y was normal in all four spores, showing that its transport to the vacuole was unaffected (results not shown). Therefore, the defect in maturation does not represent a general transport defect. These data are consistent with a role for Gpi10p in GPI-anchor synthesis.

Pig-B mutant cells are defective for the addition of the third mannose to the intermediate structure Man-(EthN-P)Man-GlcN-PI [32]. To test if *GPI10* also mediates the addition of the third mannose to the GPI intermediate, we performed lipid labelling experiments *in vivo*. Wild-type yeast cells incorporate little exogenously added mannose. We therefore used cells carrying the *pmi40* mutation, which is a temperature-sensitive allele of phosphomannose isomerase. Upon shift to non-permissive temperature, these cells incorporate exogenously added mannose with high efficiency. The *pmi40* mutation was introduced into a *gpi10*Δ::KanMX strain containing pGal-GPI10 (RH3999). A *pmi40 GPI10* strain derived from the same cross was used as control (RH3998). Cells were grown to saturation in YPG and shifted to SDCU medium (5% glucose) for 16 h. Cells were preincubated for 10 min in the presence or absence of the GPI-synthesis inhibitor, YW3548, which causes the accumulation of Man_2-GlcN-(acyl)PI [24]. After labelling for 45 min with [^{14}C]mannose, total lipids were extracted and analysed by TLC (Figure 4). In wild-type cells, one can see some labelling of the complete GPI precursor and mannosylated sphingolipids. When the inhibitor YW3548 was added, a strong accumulation of a band comigrating with Man_2-GlcN-(acyl)PI occurred [24]. In Gpi10p-depleted cells, Man_2-GlcN-(acyl)PI accumulated even in the absence of YW3548. In addition, a more polar lipid (indicated with an arrow in Figure 4) was strongly labelled, which was sensitive to treatment with GPI phospholipase D, but resistant to digestion with PI phospholipase C, suggesting that the lipid had an acylated inositol ring (results not shown). The appearance of this lipid is strongly diminished by YW3548. From the novel lipid and from the Man_2-GlcN-(acyl)PI, hydrophilic fragments were prepared by nitrous acid deamination and borohydride reduction and were analysed by Bio-Gel P4 chromatography. Furthermore, neutral glycans were prepared by treatment with HF, followed by nitrous acid deamination and borohydride reduction, and were analysed by Dionex chromatography (Table 2). In addition, the deaminated, reduced fragment derived from the novel lipid was subjected to treatment with Jack bean mannosidase followed by analysis by Dionex chromatography (Table 2). The novel lipid is sensitive to treatment with HF, contains the core glycan Man_2-anhydromannitol and elutes from a Bio-Gel P4 sizing column at 7.3 glucose units (GU). On the same column, a deaminated, reduced fragment derived from *T. brucei* lipid C (EthN-P-Man_3-GlcN-PI) eluted at 7.8 GU. Jack bean mannosidase treatment resulted in a shift in the elution profile from 7.3 to 7.0 GU and the release of a fragment eluting at 0.9 GU (mannose). These findings are consistent with a structure Man-(EthN-P)Man-GlcN-(acyl)-PI. One additional ethanolamine phosphate leads to an apparent size increase of about 4 GU as previously demonstrated by comparing the elution position of EthN-P-Man_3-GlcN-PI and Man_3-GlcN-PI [28].

GPI10 overexpression results in partial resistance to YW3548

As YW3548 prevents the addition of the third mannose to the GPI intermediate Man_2-GlcN-(acyl)PI and our data suggest that

Table 2 Analysis of the lipid intermediates accumulated upon depletion of Gpi10p

RH3999 cells were metabolically labelled with [^{14}C]Man, and total lipids were extracted and analysed by TLC. The bands corresponding to Man_2-GlcN-(acyl)PI, the novel lipid (Figure 4, arrow) and of lipid C [EthN-P-Man_3-GlcN-(acyl)PI] from *Trypanosoma brucei* were eluted from the TLC material by organic solvents. Hydrophilic fragments were generated by nitrous acid deamination and reduction, whereas neutral core glycans were generated by dephosphorylation, deamination and reduction. Hydrophilic fragments and neutral core glycans were analysed by Bio-Gel P4 size-exclusion and Dionex high-pH anion-exchange chromatography (HPAEC) respectively. Partially hydrolysed glucose oligomers were used as internal standards. The elution position of the hydrophilic fragments and the neutral core glycans are indicated as glucose units (GU) and Dionex units (DU) respectively. JBAM, Jack bean α-mannosidase.

Lipid	Bio-Gel P4 Gel filtration analysis of the hydrophilic fragments (GU)	Dionex HPAEC analysis of the neutral core glycans (DU)
EthN-P-Man_3-GlcN-(acyl)PI	7.8	2.5
Man_2-GlcN-(acyl)PI	3.3	2.1
Novel lipid	7.3	2.1
Novel lipid (JBAM)	7.0/0.9	1.1

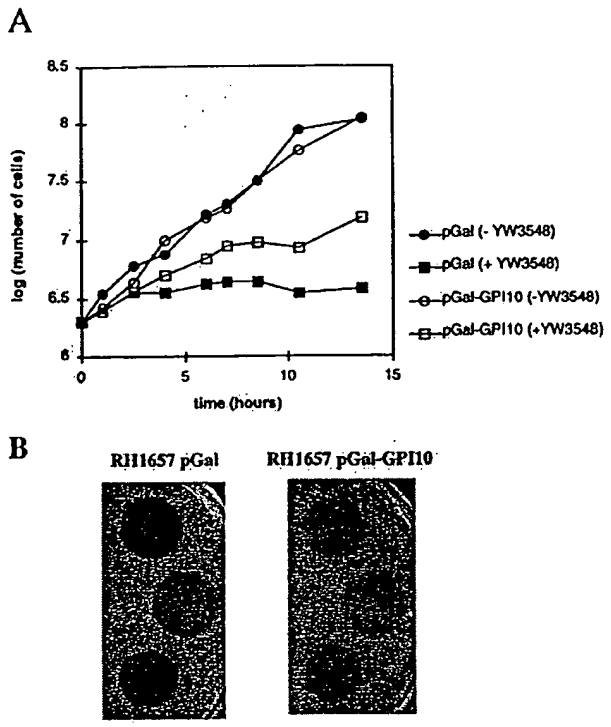


Figure 5 *GPI10* overexpression confers partial resistance to the GPI-synthesis inhibitor YW3548

(A) RH1657 cells carrying pGal or pGal-GPI10 were grown in YGal and incubated with 2 μg/ml YW3548 or methanol and growth was monitored by counting cells in a haemocytometer. (B) RH1657 cells carrying pGal or pGal-GPI10 were included in YGal plates and 4 μl of different concentrations of YW3548 were spotted on the plates (top, 100 μg/ml; middle, 10 μg/ml; and bottom, 1 μg/ml) before incubation at 30 °C.

GPI10 is required for this step, we decided to test whether overexpression of *GPI10* would lead to resistance to YW3548. To address this question, we transformed wild-type cells with either pGal-GPI10 or the empty pGal control vector. The growth of cells in YGal medium in the presence of YW3548 or methanol as control was monitored (Figure 5A). We observed that cells carrying the control vector were unable to grow in the presence of YW3548, whereas cells that expressed *GPI10* from the GAL1/10 promoter were able to overcome the effects of the compound and grow, albeit more slowly, than untreated cells. In a different assay, cells carrying either pGal-GPI10 or pGal were included in YGal plates and halo assays were performed (Figure 5B). Cells carrying the control vector gave a completely clear halo upon addition of YW3548. In contrast, overexpression of *GPI10* allowed partial growth in the presence of the compound. Whereas the halo was clear in the beginning, it became turbid after 2 days of incubation. This partial resistance suggests that YW3548 might interfere directly or indirectly with the function of Gpi10p.

DISCUSSION

In this study, we present the characterization of the *GPI10* gene from yeast. A deletion of the essential *GPI10* gene can be complemented by expression of the human *PIG-B* cDNA, suggesting that *GPI10* and *PIG-B* are functional homologues. Both genes appear to mediate the same step in GPI-anchor synthesis. Previous studies have indicated a high degree of homology between the mammalian and the yeast GPI-anchor synthesis machinery. The *gpi8* ts mutant was shown to be complemented by its human homologue [21]. Also, mammalian class E mutants that are defective in Dol-P-Man synthase can be complemented by the homologous yeast gene [33]. In this study we show that a complete deletion of *GPI10* can be complemented by its human homologue, the *PIG-B* gene.

As expected from the sequence homology and the functional complementation test, *GPI10* is involved in GPI-anchor synthesis. When *GPI10* expression was repressed, radiolabelled inositol incorporation into proteins was greatly reduced and Gas1p maturation was strongly delayed, which is consistent with a GPI-synthesis defect. All other genes that have been implicated in yeast GPI-anchor synthesis have also been found to be essential [18–22,34].

When *GPI10* expression was turned off, Man_2-GlcN-(acyl)PI as well as a novel lipid were found to accumulate. This lipid contained a Man_2-anhydromannitol core glycan after dephosphorylation, nitrous acid treatment and reduction. Its hydrophilic fragment, however, eluted from a Bio-Gel P4 column at a much later position, which could be due to an ethanolamine phosphate attached to Man_2-anhydromannitol. Furthermore, it is sensitive to treatment with Jack bean mannosidase. These findings suggest that the novel lipid may represent a GPI intermediate, most likely Man-(EthN-P)Man-GlcN-(acyl)PI. For the moment there is no formal proof that the aqueous HF-sensitive group on the mannose is ethanolamine phosphate. At this point different possibilities exist: it could represent an aberrant structure, which is only synthesized due to extensive pressure on the GPI-synthesis machinery imposed by the depletion of Gpi10p, or it could represent a natural intermediate. Most GPI anchors isolated from mature yeast proteins do not seem to carry an ethanolamine phosphate modification on the first mannose [35]. However, in animal cells, GPI molecules have been identified with this modification [1], and the mammalian PIG-B enzyme most likely acts on a GPI precursor with the same carbohydrate structure that accumulated in this study. Previous studies, which characterized the complete, radiolabelled GPI precursor in yeast before it was added to the protein, would not necessarily have detected this modification with the techniques used [27].

Previously, we suggested that YW3548 blocked the addition of the third mannose to the GPI core because when cells or membranes were treated with the compound, a Man_2-GPI intermediate with two mannoses and no ethanolamine accumulated. Here we have observed that the addition of a putative ethanolamine phosphate to the Man_2-GPI intermediate was prevented by incubation of cells with YW3548. This raises the possibility that the inhibition of the addition of the third mannose to the GPI core is indirect and is due to the inhibition of ethanolamine phosphate addition to the first mannose by YW3548. Man-(EthN-P)Man-Glc-(acyl)PI could be the natural and highly preferred substrate for Gpi10p, the putative α1,2 mannosyltransferase that adds the third mannose. In this case, the biosynthetic scheme in yeast would be PI → GlcNAc-PI → GlcN-PI → GlcN-(acyl)PI → Man-GlcN-(acyl)PI → Man-Man-GlcN-(acyl)PI → Man-(EthN-P)Man-GlcN-(acyl)PI → Man-Man-(EthN-P)Man-GlcN-(acyl)PI → (Man)Man-Man-(EthN-P)Man-GlcN-(acyl)PI → EthN-P-(Man)Man-Man-(EthN-P)Man-GlcN-(acyl)PI and YW3548 would block the addition of ethanolamine phosphate to the Man_2 intermediate. Consistent with this synthetic scheme we have preliminary data on the structure of a GPI product that accumulates in a mutant that is defective in addition of the mannose side-chain. This product

also appears to have an EthN-P on the first mannose of the core GPI structure (C. Sütterlin, P. Gerold, R. Schwarz and H. Riezman, unpublished work). This newly proposed sequence of events in the yeast GPI synthesis pathway would also be consistent with the lack of activity of YW3548 on all protozoa tested [24], because these cells do not use this modification on the first mannose [1]. Previously, we suggested that YW3548 might mimic the Man_2-GlcN-(acyl)PI substrate and compete for binding of this substrate to the α-1,2 mannosyltransferase [24]. This same argument would hold for the phosphoethanolamine transferase needed to generate this novel yeast lipid.

Overexpression of *GPI10* can confer partial resistance to YW3548, as tested by growth in the presence of YW3548. The partial resistance could be explained if Gpi10p were one of the targets of YW3548 or if the real target is the phosphoethanolamine transferase that works on the first mannose. In the latter case, Gpi10p may utilize the unmodified Man_2-GPI only poorly as a substrate. Although no mannosyltransferase activity of Gpi10p and the Pig-B protein have been shown, both proteins share significant homology with Alg9p (18% identity, 31% similarity). Alg9p is involved in the transfer of the seventh mannose in N-glycan precursor synthesis [36]. The mannose in that case is also donated by Dol-P-Man and is attached in an α-1,2 linkage. It is therefore likely that Pig-B, Gpi10p and Alg9p represent a new family of Dol-P-Man-requiring α-1,2 mannosyltransferases.

The identification of *GPI10* as a gene involved in an intermediate step of GPI-anchor synthesis improves the characterization of the yeast GPI-synthesis machinery. Functional complementation could also allow the cloning of the protozoan enzymes which, at least in their native environments, have been shown not to be sensitive to YW3548 [24]. Expression of these genes in yeast may make it possible to screen for compounds that specifically interfere with the addition of the third mannose residue in fungi or protozoa.

This work was supported by grants from the Swiss National Science Foundation, the Canton of Basel-Stadt and the Swiss Federal Office for Education and Science to H.R., and by a grant of HFSPO to H.R., R.T.S. and T.K. M.J.M. was supported by grants UE (BIO4-CT95-0080) and CICYT (BIO96-2461-CE). M.V.E. thanks FEBS for a short term fellowship. We thank a very insightful reviewer and M. Schönbächler and M. Mishkind for critical reading of the manuscript. We are grateful to Keiko Kinoshita for technical assistance. R.T.S. was supported by the Deutsche Forschungsgemeinschaft, Hessische Ministerium für Kultur und Wissenschaft, Fonds der Chemischen Industrie and Stiftung P. E. Kempkes.

REFERENCES

1 McConville, M. J. and Ferguson, M. A. J. (1993) Biochem. J. **294**, 305–324
2 Englund, P. T. (1993) Annu. Rev. Biochem. **62**, 121–138
3 Takeda, J. and Kinoshita, T. (1995) Trends Biochem. Sci. **20**, 367–371
4 Udenfriend, S. (1995) Annu. Rev. Biochem. **64**, 563–591
5 Maxwell, S. E., Ramalingam, S., Gerber, L. D., Brink, L. and Udenfriend, S. (1995) J. Biol. Chem. **270**, 19576–19582
6 Riezman, H. and Conzelmann, A. (1998) Handbook for Proteolytic Enzymes (Barret, A. J., Rawlings, N. D. and Woessner, J. F., eds.), Academic Press, London, in the press
7 Nuoffer, C., Horvath, A. and Riezman, H. (1993) J. Biol. Chem. **268**, 10558–10563
8 Kinoshita, T. and Takeda, J. (1994) Parasitol. Today **10**, 139–143
9 Stevens, V. L., Yhang, H. and Harrenman, M. (1996) Biochem. J. **313**, 253–258
10 Mohney, R. P., Knez, J. J., Ravi, L., Sevlever, d., Rosenberry, T. L., Hirose, S. and Medof, M. E. (1994) J. Biol. Chem. **269**, 6536–6542
11 Nakamura, N., Inoue, N., Watanabe, R., Takahashi, M., Takeda, J., Stevens, V. L. and Kinoshita, T. (1997) J. Biol. Chem. **272**, 15834–15840
12 Sugiyama, E., DeGasperi, R., Urakaze, M., Chang, M., Thomas, L. J., Hyman, R., Warren, C. D. and Yeh, E. T. H. (1991) J. Biol. Chem. **266**, 12119–12122
13 Takahashi, M., Inoue, N., Ohishi, K., Maeda, Y., Nakamura, N., Endo, Y., Takeda, J. and Kinoshita, T. (1996) EMBO J. **15**, 4254–4261
14 Miyata, T., Takeda, J., Iida, Y., Yamada, N., Inoue, N., Takahashi, M., Maeda, K., Kitani, T. and Kinoshita, T. (1993) Science **259**, 1318–1320
15 Inoue, I., Watanabe, R., Takeda, J. and Kinoshita, T. (1996) Biochem. Biophys. Res. Commun. **226**, 193–199
16 Inoue, N., Kinoshita, T., Orii, T. and Takeda, J. (1993) J. Biol. Chem. **268**, 6882–6885
17 Kamitani, T., Chang, H., Rollins, C., Waneck, G. L. and Yeh, E. T. H. (1993) J. Biol. Chem. **268**, 20733–20736
18 Leidich, S. D., Kostova, Z., Latek, R. R., Costello, L. C., Drapp, D. A., Gray, W., Fassler, J. S. and Orlean, P. (1995) J. Biol. Chem. **270**, 13029–13035
19 Schönbächler, M., Horvath, A., Fassler, J. and Riezman, H. (1995) EMBO J. **14**, 1637–1645
20 Hamburger, D., Egerton, M. and Riezman, H. (1995) J. Cell Biol. **129**, 629–639
21 Benghezal, M., Benachour, A., Rusconi, S., Aebi, M. and Conzelmann, A. (1996) EMBO J. **15**, 6575–6583
22 Benghezal, M., Lipke, P. N. and Conzelmann, A. (1995) J. Cell Biol. **130**, 1333–1344
23 Menon, A. K., Mayor, S. and Schwarz, R. T. (1990) EMBO J, **9**, 4249–4258
24 Sütterlin, C., Horvath, A., Gerold, P., Schwarz, R. T., Wang, Y., Dreyfuss, M. and Riezman, H. (1997) EMBO J. **16**, 6374–6383
25 Wach, A., Brachat, A., Pöhlmann, R. and Philippsen, P. (1994) Yeast **10**, 1793–1808
26 Horvath, A., Sütterlin, C., Manning-Krieg, U., Movva, N. R. and Riezman, H. (1994) EMBO J. **13**, 3687–3695
27 Sipos, G., Puoti, A. and Conzelmann, A. (1994) EMBO J. **13**, 2789–2796
28 Gerold, P., Striepen, B., Reitter, B., Geyer, H., Geyer, R., Reinwald, E., Risse, H.-J. and Schwarz, R. T. (1996) J. Mol. Biol. **261**, 181–194
29 Ferguson, M. A. J. (1991) Lipid Modification of Proteins: A Practical Approach (Hooper, N. M. and Turner, A. J., eds.), IRL Press, Oxford
30 Nuoffer, C., Jenö, P., Conzelmann, A. and Riezman, H. (1991) Mol. Cell. Biol. **11**, 27–37
31 Doering, T. L. and Schekman, R. (1996) EMBO J. **15**, 182–191
32 Puoti, A., Desponds, C., Fankhauser, C. and Conzelmann, A. (1991) J. Biol. Chem. **268**, 21051–21059
33 DeGasperi, R., Thomas, L. J., Sugiiyama, E., Chang, H. M., Beck, P. J., Orlean, P., Albright, C., Sambrook, J. F., Warren, C. D. and Yeh, E. T. H. (1990) Science **250**, 988–991
34 Leidich, S. D., Drapp, D. A. and Orlean, P. (1994) J. Biol. Chem. **269**, 10193–10196
35 Fankhauser, C., Homans, S. M., Thomas-Oates, J. E., McConville, M. J., Desponds, C., Conzelmann, A. and Ferguson, M. A. J. (1993) J. Biol. Chem. **268**, 26365–26374
36 Burda, P., te Heesen, S., Brachat, A., Wach, A., Dusterhof, A. and Aebi, M. (1996) Proc. Natl. Acad. Sci. U.S.A. **93**, 7160–7165

Received 11 September 1997/18 February 1998; accepted 19 February 1998

In re: Meyers
Appl. No. 09/464,039
Filed December 15, 1999

APPENDIX E

The EMBO Journal Vol.16 No.14 pp.4340–4351, 1997

Human and *Xenopus* cDNAs encoding budding yeast Cdc7-related kinases: *in vitro* phosphorylation of MCM subunits by a putative human homologue of Cdc7

Noriko Sato, Ken-ichi Arai and Hisao Masai[1]

Department of Molecular and Developmental Biology, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108, Japan

[1]Corresponding author

***Saccharomyces cerevisiae* Cdc7 kinase is essential for initiation of DNA replication, and Hsk1, a related kinase of *Schizosaccharomyces pombe*, is also required for DNA replication of fission yeast cells. We report here cDNAs encoding Cdc7-related kinases from human and *Xenopus* (huCdc7 and xeCdc7, respectively). The cloned cDNA for huCdc7 contains an open reading frame consisting of 574 amino acids with a predicted molecular weight of 63 847 that possesses overall amino acid identity of 32% (54% including similar residues) to Cdc7 and Hsk1. huCDC7 is transcribed in the various tissues examined, but most abundantly in testis. Three transcripts of 4.4, 3.5 and 2.4 kb in length are detected. The 3.5 kb transcript is the most predominant and is expressed in all the tissues examined. A cDNA containing a 91 nucleotide insertion at the N-terminal region of huCDC7 is also detected, suggesting the presence of multiple splicing variants. The huCdc7 protein is expressed at a constant level during the mitotic cell cycle and is localized primarily in nuclei in interphase and distributed diffusibly in cytoplasm in the mitotic phase. The wild-type huCdc7 protein expressed in COS7 cells phosphorylates MCM2 and MCM3 proteins *in vitro*, suggesting that huCdc7 may regulate processes of DNA replication by modulating MCM functions.**

Keywords: CDC7/cell cycle/DNA replication/ MCM proteins/serine-threonine kinase

Introduction

Initiation of chromosomal replication in eukaryotes is tightly regulated during the cell cycle. Genetic and biochemical studies in the yeast *Saccharomyces cerevisiae* have yielded considerable information on molecular interactions between replication machinery and cell cycle regulators (Coverley and Laskey, 1994; Huberman, 1995; Kearsey *et al.*, 1996). The replication origins of *S.cerevisiae*, which are composed of the conserved 11 bp 'A' or core sequence and auxiliary 'B' elements, exist as nucleoprotein structures which involve origin recognition complex (ORC), MCM, Cdc6 and probably other proteins (Campbell and Newlon, 1991; Walker *et al.*, 1991; Bell and Stillman, 1992; Diffley and Cocker, 1992; Marahrens and Stillman, 1992; Bell *et al.*, 1993; Diffley *et al.*, 1994; Rao *et al.*, 1994; Rowley *et al.*, 1994, 1995; Theis and Newlon, 1994; Fox *et al.*, 1995; Liang *et al.*, 1995; Loo *et al.*, 1995; Rao and Stillman, 1995). It was reported previously that these protein–DNA complexes may alternate between two distinct states during the cell cycle; one that exists prior to the S phase and the other that is detected during the G_2 and M phases. The former pre-replicative complex contains ORC and Cdc6 as well as MCM licensing factor, while the latter post-replicative complex may contain only ORC (Liang *et al.*, 1995; Cocker *et al.*, 1996; Donovan and Diffley, 1996). The pre-replicative complex may be activated by regulatory molecules which trigger the initiation of S phase. Genetic study of *S.cerevisiae* has implicated serine-threonine kinases in this step, among which the Cdc7–Dbf4 kinase complex may turn on the ultimate 'START' signal for the S phase (Kitada *et al.*, 1992; Jackson *et al.*, 1993; Dowell *et al.*, 1994; Sclafani and Jackson, 1994; Bell, 1995). In the presence of active Cdc7 kinase, the S phase can be completed in the absence of protein synthesis (Hartwell, 1974). Cdc7, whose kinase activity peaks at the G_1/S boundary, activates DNA replication machinery in conjunction with Dbf4, which not only stimulates its kinase activity but may also tether Cdc7 at the origins of replication (Jackson *et al.*, 1993; Yoon *et al.*, 1993; Dowell *et al.*, 1994).

The structures of DNA replication origins and modes of their activation have been elusive in higher eukaryotes (Hamlin and Dijkwel, 1995). However, identification of genes related to ORC and MCM components in *Xenopus*, *Drosophila* and mammals has strongly indicated that the basic components required for initiation of chromosomal replication may be conserved in higher eukaryotes (Thommes *et al.*, 1992; Hu *et al.*, 1993; Chong *et al.*, 1995; Ehrenhofer-Murray *et al.*, 1995; Gavin *et al.*, 1995; Gossen *et al.*, 1995; Kimura *et al.*, 1995; Kubota *et al.*, 1995; Treisman *et al.*, 1995; Carpenter *et al.*, 1996). We previously reported *hsk1*$^+$, whose product is a putative *Schizosaccharomyces pombe* homologue of Cdc7 kinase (Masai *et al.*, 1995). *hsk1*$^+$ is essential for viability of *S.pombe* cells, and analyses of DNA content and morphology of germinating spores containing *hsk1* null alleles indicated that *hsk1*$^+$ is required for DNA replication as well as for coupling of the M phase to S phase initiation. The presence of the structurally and functionally related kinases in two distantly related yeast species suggested the possibility that eukaryotic DNA replication may be regulated through a conserved mechanism which involves Cdc7-related kinases.

We report here isolation of human and *Xenopus* cDNAs encoding Cdc7-related kinases (huCdc7 and xeCdc7, respectively). We show that huCdc7 is a nuclear protein kinase expressed at a constant level throughout the cell cycle. We also report that huCdc7 expressed in COS7 cells phosphorylates MCM components *in vitro*, suggesting possible regulation of MCM functions by Cdc7-related kinase.

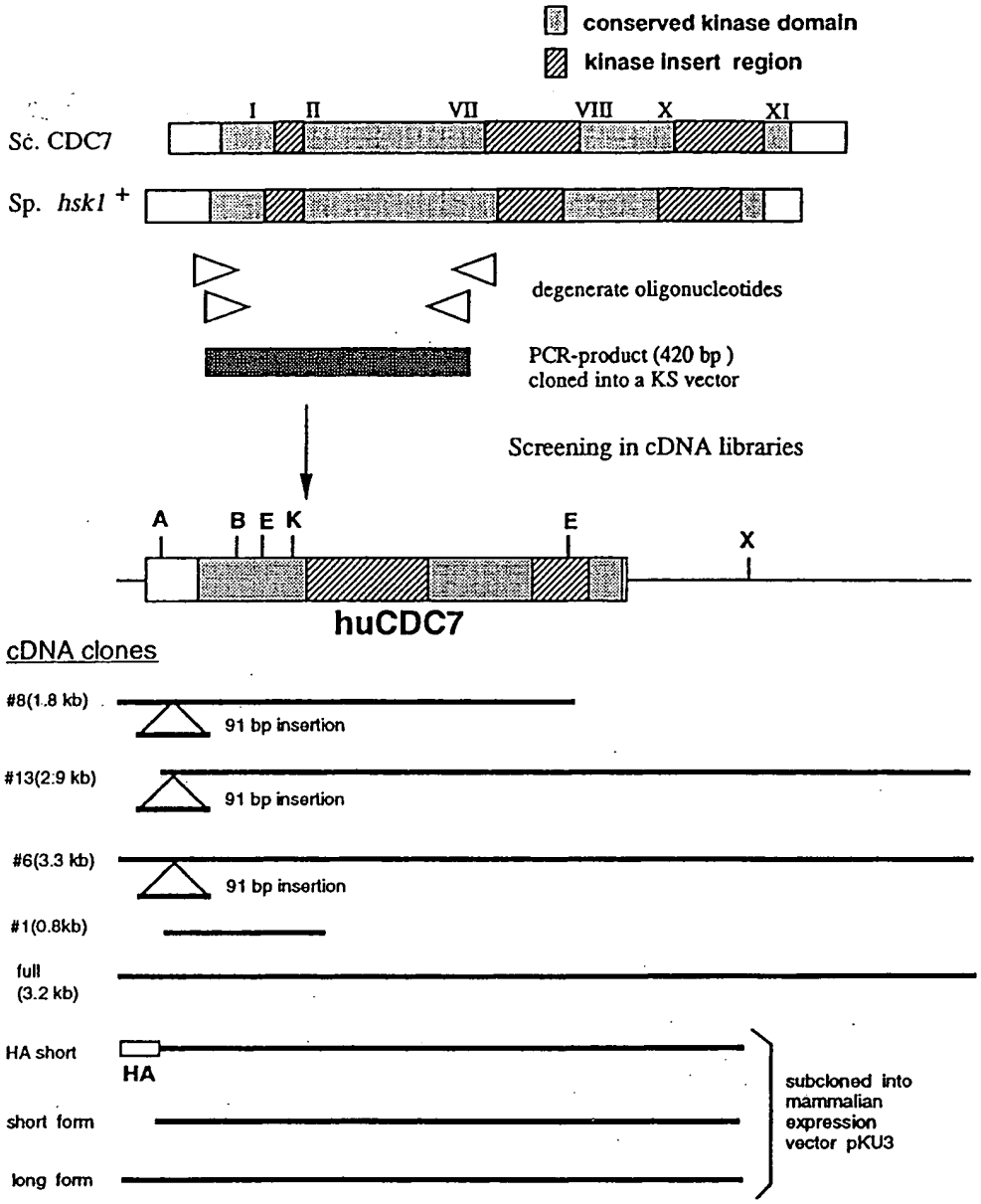


Fig. 1. Strategy for PCR cloning of CDC7-related kinases and cloned cDNAs. Two sets of degenerate primers indicated by open arrowheads were designed on the basis of the conserved amino acid sequences between the products of *S.cerevisiae* (Sc.) *CDC7* and *S.pombe* (Sp.) *hsk1*$^+$. Shaded and striped regions represent conserved kinase domains and kinase insert sequences, respectively, whereas open areas indicate N- and C-terminal tail sequences. A, B, E, K and X represent the sites for *AgeI*, *BclI*, *EcoRI*, *KasI* and *XbaI*, respectively. HA indicates a haemagglutinin epitope sequence. The thick bars indicate the segments present in the isolated cDNAs or those subcloned into vectors. Details are described in the text.

Results

Isolation of human and Xenopus cDNAs related to S.cerevisiae CDC7 and S.pombe hsk1$^+$

In order to isolate *CDC7/hsk1*$^+$-related genes from higher eukaryotes, we performed RT–PCR on mRNA isolated from murine embryonic stem (ES) cells, using degenerate oligonucleotide probes designed on the basis of amino acid sequence homology between the products of *CDC7* and *hsk1*$^+$. A PCR product ~420 bp long, generated with a set of the primers derived from the conserved amino acid sequences of kinase subdomains I and VII, was isolated and used as a template for the second nested PCR with primers derived from the conserved domains I and VI. The second PCR gave rise to one major band of expected size (~350 bp). Therefore, the 420 bp long PCR product from the first PCR was subcloned into the T vector (Marchuk *et al.*, 1991) and its sequence was determined.

Sequencing of this DNA fragment revealed the presence of a reading frame that showed substantial homology (52 identical amino acids from 138 residues) to both Cdc7 and Hsk1. Screening of human and *Xenopus* cDNA libraries using this DNA fragment as a probe led to isolation of three human and three *Xenopus* clones, and the inserts of these positive clones were subcloned into KS vectors for further characterization (Figure 1).

The restriction mapping of the 1.8 and 2.9 kb inserts from the human clones #8 and #13 (both obtained from a fetal liver library), respectively, indicated the presence of an overlapping region between the two inserts. The nucleotide sequence of 1756 bp long #8 cDNA revealed the presence of a 334 bp 5′-untranslated region and a 1422 bp stretch of a coding region. The clone #13 cDNA did not contain 5′-untranslated region, but possessed a 1416 bp stretch of a coding region and a 1334 bp 3′-untranslated region together with poly(A) sequences. Although a 535 amino acid long coding region could be deduced by combining #8 and #13 cDNAs (#6), a start codon ATG was not found upstream of the kinase subdomain I. The N-terminal region of clone #8 was identical to the portion of the sequence of #1 (obtained from a testis library), except that #1 cDNA had a 91 bp deletion. This 91 bp deletion would give rise to an open reading frame (ORF) that contains three possible ATG codons, among which the first (from the 5′ end) is used to initiate translation of huCdc7 protein (see below). The full-length huCDC7 cDNA contains a 1722 bp long ORF, encoding a 574 amino acid long protein with a predicted mol. wt of 63 847 (Figure 2).

Xenopus clones #25 and #28, obtained from a *Xenopus* oocyte library, were identical and the nucleotide sequence of the 1467 bp insert of #25 cDNA showed a 1425 bp long ORF frame and a 42 bp 3′-untranslated region. Clone #23, obtained from another *Xenopus* oocyte library, contained an additional eight amino acids at the N-terminus including the putative ATG initiation codon, and a 1449 bp long ORF, encoding a 483 amino acid protein with a predicted mol. wt of 53 509, was deduced by combining the three clones.

Primary structures of Cdc7-related kinases from higher eukaryotes

The predicted primary structures of human and *Xenopus* Cdc7-related kinase indicate that their kinase domains are highly homologous to those of Cdc7 and Hsk1 (Figure 3A). When confined to the kinase conserved domains, 44% identity at the amino acid level (62% including similar residues) is detected between Cdc7 and huCdc7 or xeCdc7, whereas 80% identity (90% including similar residues) is detected between human and frog. Amino acid sequence alignments of the kinase domains of Cdc7-related kinases were calculated together with other representative serine/threonine kinases, using the ClustalW program, and a hypothetical phylogenetic tree was drawn. In the phylogenetic tree, Cdc7-related kinases were classified into a subfamily distinct from other kinases, including CDK, CKII or MAPK (Figure 3B), indicating that they are members of a unique 'Cdc7' kinase family.

Cdc7 and Hsk1 are characterized by the presence of three 'kinase-insert' sequences between the kinase domains I and II, VII and VIII, and X and XI, designated as kinase insert I, II and III, respectively. huCdc7 and xeCdc7 also contain two amino acid insertions at the same locations (corresponding to kinase inserts II and III), although the presence of kinase insert I was not obvious in the human and frog clones. The lengths and sequences of the kinase inserts are not conserved between yeasts and higher eukaryotes as they are not between the two yeast species, although weak homology is identified between the human

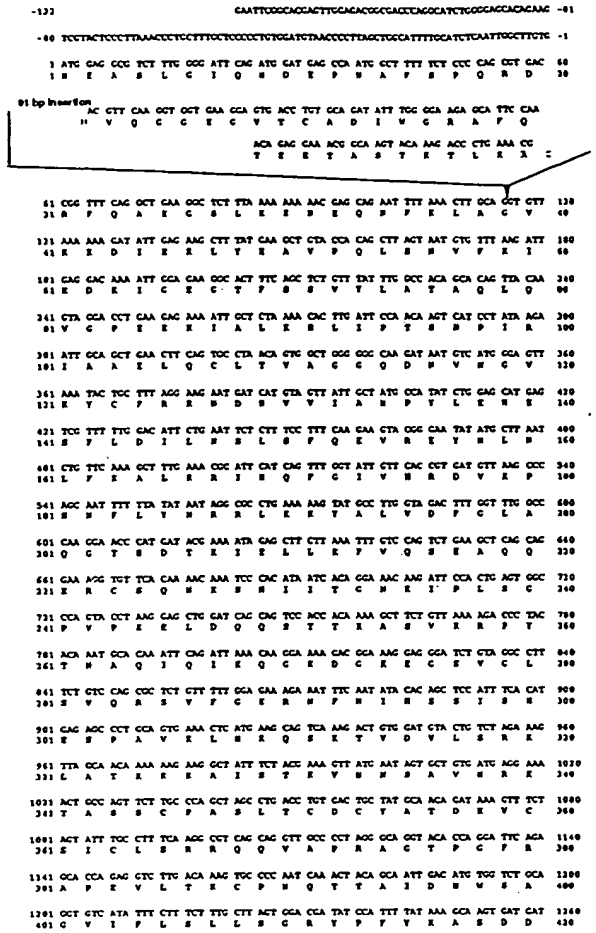


Fig. 2. Nucleotide sequences of a full-length huCDC7 cDNA and predicted amino acid sequences. The A residue of the putative first methionine codon is taken as +1. The position and nucleotide sequence of the 91 bp insertion is also shown. The poly(A) addition signal is double-underlined.

A

```
                                                                                  <I>
1 huCdc7 MEASLGIQMDEPMAF SPQRDRFQAEGSLKK NEQNFKL-------- ---AGVKKDIEKLYE AVPQLSNVFKIEDKI GEGTFSSVYLATAQL   79
2 xeCdc7 --------------- ------------MSS GD-N----------- ---SGVAKEIEKLYA AVPQLHNIFYIKSKI GEGTFSSVYFAIGRL   48
3 ScCdc7 --------------- --------MTSKTKN IDDIP---------- ---PEIKEEMIQLYH DLPGIENEYKLIDKI GEGTFSSVYKAKDIT   54
4 SpCdc7 -MAEAHITLSPKVTH EQQTDIDSECEITEV DDENVNENKSQEMIQ DIPARDREEIENITR TFVELQENYRLIEKI GEGTFSSVYKAEDLH   89
                                                              .. .      .  . .  ** ********* *

         kinase insert 1                 <II>            <III>           <IV>                <V>
1 huCdc7 Q-------------- ----VG--------- --PEEKIALKHLIPT SHPIRIAAELQCLTV AGGQDNVMGVKYCFR KNDHVVIAMPYLEHE  140
2 xeCdc7 R-------------- ----SG--------- --EDAKFALKHLIPT SHPTRIAAELQCLSV AGGEDNVMGVKYCFR NKDHVVIVMPYLEHE  109
3 ScCdc7 C-------------- ----KITKKFASHFW NYGSNYVALKKIYVT SSPQRIYNELNLLYI MTGSSRVAPLCDAKR VRDQVIAVLPYYPHE  126
4 SpCdc7 YGRYINDWDIQSEVL KESSFGKEKIPVNED SRKPKYVAIKKIYAT SSPARIYNELEILYL LRGSSVIAPLITALR NEDQVLVVLPYYEHT  179
                                     *.*..  * * * **  **  * .   *   .  .  . *   *.*.  .** *

                                     <VI>                      <VII>           kinase insert 2
1 huCdc7 SFLDILNSLSFQEVR EYMLNLFKALKRIHQ FGIVHRDVKPSNFLY NRRLKKYALVDFGLA QGTHDTKIELLKFVQ SEAQQERCSQNKSHI  230
2 xeCdc7 CFADILHSLSFEETK EYMFNLLKALRHIHS FGIVHRDVKPSNFLF NRSLKKFALVDFGLA QGTSDTKIDLLKVLQ ---------------  184
3 ScCdc7 KFRTFYRDLPIKGIK KYIWELLRALKFVHS KGIIHRDIKPTNFLF NLELGRGVLVDFGLA EAQMDYKSMISSQND ---------------  201
4 SpCdc7 DFRQYYSTFSYRDMS IYFRCLFQAMQQTQT LGIIHRDIKPSNFLF DVRTKHGVLVDFGLA ER------------- ---------------  241
          *            *   * .*..  .   **.***.**.***.    .  ******* .

1 huCdc7 ITGNKIPLSGPVPKE LDQQSTTKASVKRPY TNAQIQIKQGKDGKE GSVGLSVQRSVFGER NFNIHSSISHESPAV KLMKQSKTVDVLSRK  320
2 xeCdc7 ------------PKK QD------------- --------------- GLVGSSTQRSVFGER NFNVHSAVTIDNTTL KAAKPSKTIDVTTRK  234
3 ScCdc7 --------------- YD------------- --------------- -NYANTNHDGGYSMR N---HEQFCPCIMRN QYSPNSHNQTPPMVT  244
4 SpCdc7 --------------- YD------------- --------------- ------------GRQ Q--SHSCRCTNSNAA ELAHDFSIAQ-----  269
                         *                                               * ..

                                                                 <VIII>           <IX>
1 huCdc7 LATKKKAISTKVMNS AVMRKTASSCPASLT CDCYATDKVCSICLS RRQQVAPRAGTPGFR APEVLTKCPNQTTAI DMWSAGVIFLSLLSG  410
2 xeCdc7 LATR-KTVSTKSTSS AVPKKAASTCQTSLT CDCYAKDQVCNICLA RTRQVAPRAGTPGFR APEALTKCPHQTTAI DMWSAGIIFLSLLSG  323
3 ScCdc7 IQNG-KVVHLNNVNG VDLTKG--------- ---YPKNE------T RRIKRANRAGTRGFR APEVLMKCGAQSTKI DIWSVGVILLSLLGR  315
4 SpCdc7 --------ETSLG-- --------------- ---YIKND------T RPSKRANRAGTRGFR APEVLFKCSSQSPKV DIWSAGVILLSFLTK  325
                                                   *  . * **** *** *** * **  *.  . *.**.*.* ** *

              <X>                 kinase insert 3
1 huCdc7 RYPFYKASDDLTALA QIMTIRGSRETIQAA KTFGKSILCSKEVP- ---------AQDLRK LCERLRG---M---- DSSTPKLTSDIQGHA  483
2 xeCdc7 RYHFFNAADDMNALA QIMTIRGSKETIQAS KCFGKSVLCSKELP- ---------SKDLRT LCEGLRSAIVLPNGN QHDIQKQRAALQMRI  403
3 ScCdc7 RFPMFQSLDDADSLL ELCTIFGWKELRKCA ALHG----LGFEASG LIWDKPNGYSNGLKE FVYDLLNKECT---- IGTFPEYSVAFETFG  397
4 SpCdc7 RFPMENSKDDVDALM EIACIFGKSEMRQCA ALHG----CTFETN- --------------- --------------- VSTLTEKRVNFR---  377
         *.  . . **  .*  .. .* *  *  ...     *        *

                                                                  <XI>
1 huCdc7 SHQPAISEKTDHKAS CLVQTPPGQYSGNSF KKGDSNSCEHCFDEY -NTNLEGWNEVPDEA YDLLDKLLDLNPASR ITAEEALLHPFFKDM  572
2 xeCdc7 MENQ----------- DGWFLPESPDITPDS PAVVRSSCVSTSDNM EQSNHNGWDRVPNEA YHLLDRLLDMNPATR ITAEEALIHPLFKNM  482
3 ScCdc7 FLQ------------ --QELHDRMSIEPQL PDPKTNMDAVDAYEL KKYQEEIWSD-HYWC FQVLEQCFEMDPQKR SSAEDLLKTPFFNEL  472
4 SpCdc7 --------------- --------------- KLILWASCGSASIYK EKLRHKPSQE-ERLC LDFLEKCLELDCNKR ISAEEALDHDFLYLD  436
                                                               .   *..  ..   *  .**. *

1 huCdc7 SL------------- --------------- --------------- --------------- -----------    574
2 xeCdc7 R-------------- --------------- --------------- --------------- -----------    483
3 ScCdc7 NENTYLLD------- --GES---TDEDD-- -----VVSSSEADLL DKDVLLISE------ -----------    507
4 SpCdc7 NLAYEKKDDDTAFDN SFGETSFEKDEDLTA KHLSHILDFKEQEET DEPTSLSKRKRSIDE ILPNDALQDGA    507
```

B

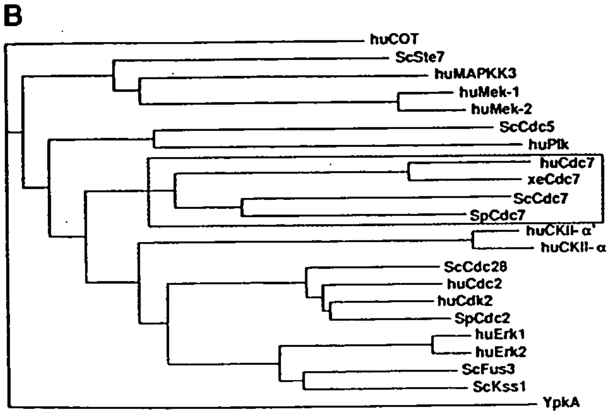


Fig. 3. Comparison of CDC7-related kinases from various organisms. **(A)** The predicted amino acid sequences of four CDC7-related kinases are aligned for maximum homology. Amino acid numbers are shown on the left. The roman figures above the sequences indicate the locations of the conserved kinase domains as previously proposed (Hanks *et al.*, 1988), which are boxed in grey. The asterisks indicate the amino acid residues conserved in all four proteins. **(B)** A hypothetical phylogenetic tree. Amino acid sequence alignments were calculated using the ClustalW program, and a phylogenetic tree was drawn by the SINCA program (Zuckerlandl and Pauling and neighbour-joining methods). huCdc7, xeCdc7, ScCdc7 and SpCdc7 indicate the Cdc7-related kinases from human, *Xenopus*, *S.cerevisiae* and *S.pombe*, respectively. COT, *cot* oncogene product; YpkA, enterobacterial protein kinase.

and *Xenopus* proteins. The kinase insert II of huCdc7 or xeCdc7 is 163 or 108 amino acid long, respectively, and 57 amino acids are identical, with 11 additional similar residues. The kinase insert III of huCdc7 or xeCdc7 is 98 or 95 amino acids long, respectively, of which 34 amino acids are identical and 14 additional amino acids are similar.

Another feature of Cdc7 and Hsk1 is the presence of

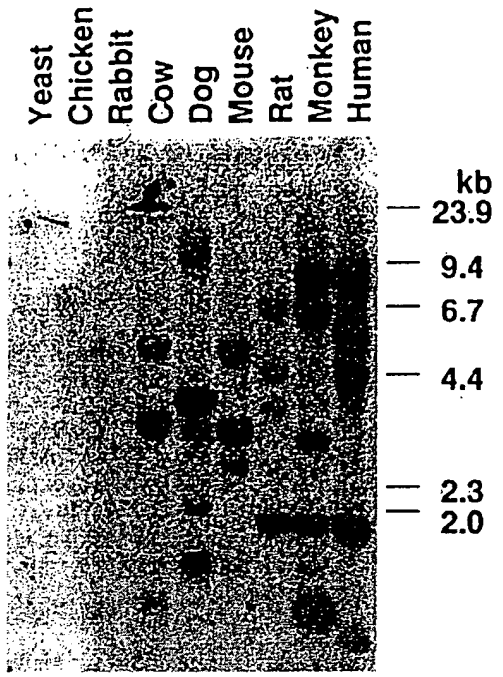


Fig. 4. Genomic Southern analysis of *CDC7*-related genes in various species. Genomic DNAs were digested with *Eco*RI. Membrane was hybridized using huCDC7 cDNA (probe A, nucleotide residues −132 to +1535; see the legend to Figure 5) as a probe, and washed in 0.1× SSC and 0.1% SDS at 50°C.

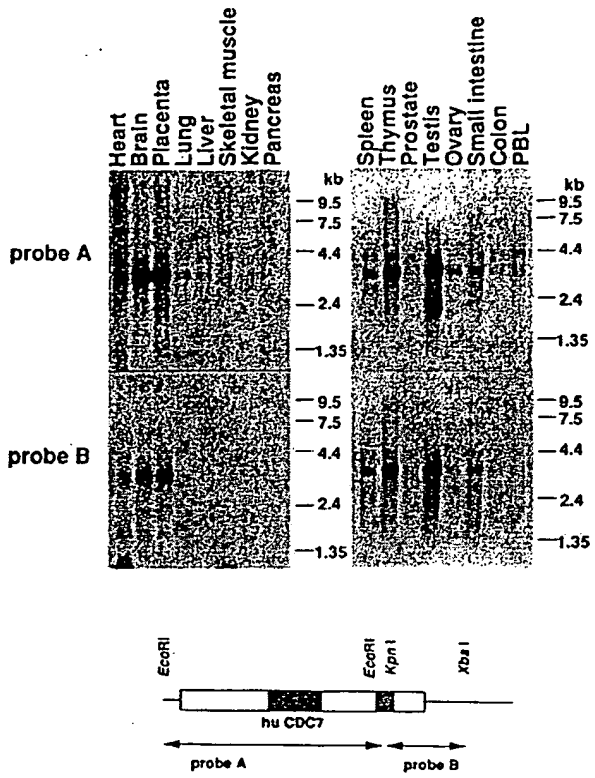


Fig. 5. Expression of huCDC7 mRNA in various human tissues. Membranes were hybridized with probe A (upper) or probe B (lower). Probes A and B include the huCDC7 cDNA fragment from nucleotide residue −132 to +1535 and that from +1611 to +2202, respectively. The shaded regions indicate the kinase insert sequences.

the C-terminal regions which are rich in acidic residues. Two-hybrid assays indicated that Dbf4 protein could interact with this C-terminal tail of Cdc7 protein in budding yeast (Jackson *et al.*, 1993). The C-terminal regions of both Cdc7 and Hsk1 are essential for the functions of these two kinases (our unpublished data). Unexpectedly, huCdc7 and xeCdc7 did not contain similar C-terminal tails (see discussion below).

Genomic Southern and Northern analysis

Southern analyses of genomic DNA from various eukaryotic species using two *Eco*RI fragments (0.7 and 1.1 kb) derived from huCDC7 cDNA containing the amino acids 1–511 as a probe (probe A) indicated that they hybridized with DNA fragments of human, monkey, rat, mouse, dog, cow, rabbit and chicken under a stringent washing condition (0.1× SSC and 0.1% SDS at 50°C), but that they did not hybridize with those of budding yeast even under a relaxed washing condition (2× SSC and 0.1% SDS at 42°C) (Figure 4).

Northern analyses of mRNA from various tissues (Figure 5) showed that huCDC7 transcripts, which were detected in most tissues examined, are expressed at a high level in testis and at a moderate level in thymus, spleen, placenta, brain and heart. Three transcripts of 4.4, 3.5 and 2.4 kb in length were detected, among which the 3.5 kb transcript was ubiquitously detected. The 2.4 kb transcript was seen only in testis, and did not hybridize with a probe specific to the C-terminal region of huCDC7 cDNA (*Kpn*I–*Xba*I fragment containing the amino acids 538–574; probe B), indicating that it is an alternatively spliced form lacking the C-terminal coding region. Transcription of a mouse Cdc7-related kinase also exhibited similar tissue distribution, with the highest expression in testis (our unpublished data).

The 3.5 kb transcript was predominant and ubiquitously detected in various human cell lines such as HL60, K562, MOLT4, Raji, SW480 and HeLa (S3) (data not shown). The level of huCDC7 transcription was similar between these cancer cell lines. The 4.4 kb transcript was also detected in most cell lines, albeit at a much lower level.

Identification of endogenous huCdc7 protein

Endogenous huCdc7 was detected by specific antibodies, which were raised against GST fusion proteins containing segments of huCdc7 protein. The mouse monoclonal antibody 4A8 specifically recognized a single protein which was expressed by the huCdc7-expressing plasmid in COS7 cells and migrated with an apparent mol. wt of 68 kDa (Figure 6A). huCdc7 protein was immunoprecipitated by the 4A8 antibody from the cell lysates prepared from a factor-dependent myeloid leukaemia cell line, TF-1, and was identified by Western blot analysis with the rabbit polyclonal antibody #1 (Figure 6B). In order to determine which of the three possible ATG codons is utilized to initiate translation of the huCdc7 coding frame, we expressed both long and short forms of huCdc7 (pKU-long-huCdc7 and pKU-short-huCdc7 initiated from the first and third ATG, respectively; Figure 1) and compared their migration on a gel with that of the endogenous

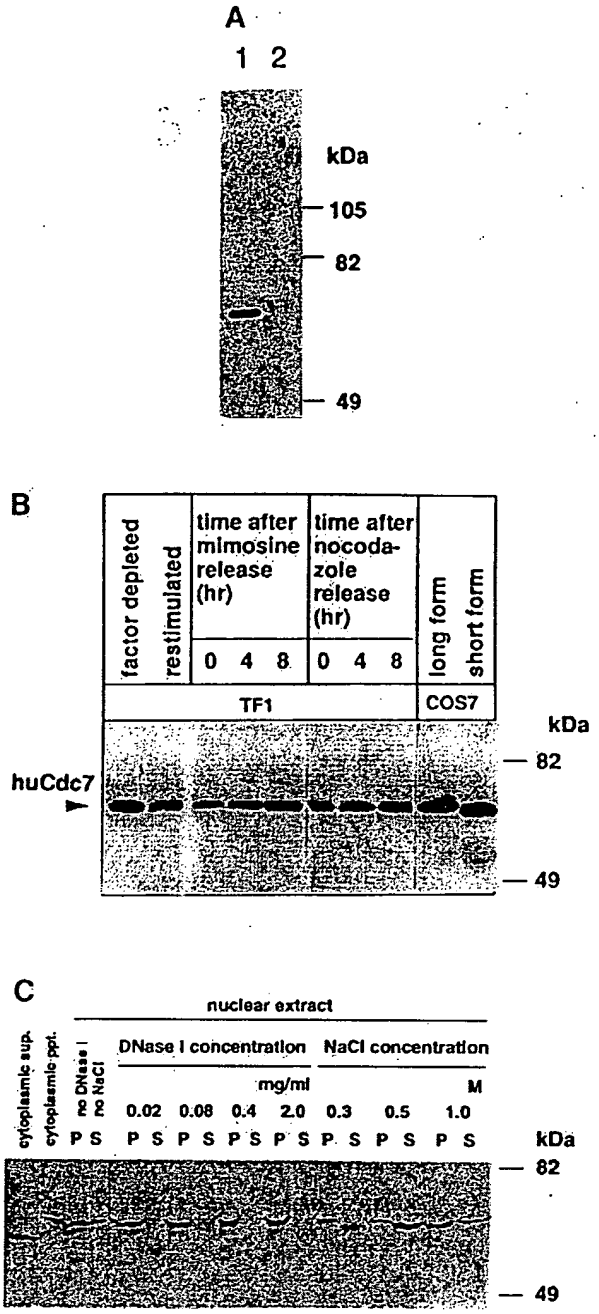


Fig. 6. Expression of huCdc7 protein in tissue culture cells. (A) Specificity of the monoclonal antibody, 4A8. Nuclear extracts prepared from COS7 cells transfected with pKU-long-huCdc7 (lane 1) or with the vector (lane 2) were separated on SDS–PAGE and blotted in 1 μg/ml monoclonal antibody 4A8. (B) Expression of huCdc7 protein at different cell cycle stages. The whole cell lysates were prepared from 10^7 TF-1 cells or from 1.5×10^6 COS7 cells which had been transfected with 1 μg of long or short form of huCDC7 expression vector and had been harvested 2 days after transfection. huCdc7 protein was immunoprecipiated with the monoclonal antibody 4A8. (C) Salt and DNase I extraction of nuclear huCdc7 protein. K562 cells were extracted with hypotonic buffer and nuclei were extracted with buffer containing DNase I or NaCl at the concentrations indicated. The supernatant (S) and pellet (P) were separated by centrifugation. Immunoprecipitates (B) and protein fractions (C) were separated on an 8% SDS–PAGE, transferred onto Immobilon paper and were blotted with anti-huCdc7 polyclonal antibody #1.

protein. The endogenous huCdc7 protein co-migrated with the long form, indicating that the first ATG initiates translation of huCdc. (Figure 6B).

The amount of huCdc7 stayed relatively constant at various phases of the mitotic cell cycle in TF-1 cells synchronized by mimosine or nocodazole or after depletion of a growth factor, human granulocyte–macrophage colony-stimulating factor (GM-CSF) (Figure 6B).

Nuclear localization of huCdc7

huCdc7 is localized mainly in nuclei in the interphase and is present diffusibly in the cytoplasm in the mitotic phase of HeLa cells, as indicated by immunofluorescence analyses using the monoclonal antibody 4A8 (data not shown). In order to examine whether the nuclear localization of huCdc7 changes at the G_1/S transition, HeLa cells were synchronized at the G_1/S boundary by double thymidine block and the S phase cells were obtained at 4 h after the release into the cell cycle. In this experiment, soluble proteins and phospholipids were extracted with Triton X-100 before fixation (Fey *et al.*, 1984). Cells were stained by anti-bromodeoxyuridine (BrdU), anti-proliferating cell nuclear antigen (PCNA), anti-mouse CDC21 (MCM4) and anti-huCdc7 (4A8) antibodies. BrdU was incorporated into 100% of the S phase cells, indicating that DNA was being synthesized synchronously (Figure 7B). PCNA accumulated in nuclei as the S phase progressed (Figure 7D). On the other hand, MCM protein was localized in nuclei before DNA synthesis, and disappeared from nuclei during the S phase, as expected from its licensing function required for 'once and only once' replication in the S phase (Figure 7E and F) (Blow and Laskey, 1988; Blow, 1993; Kimura *et al.*, 1994; Chong *et al.*, 1995; Kubota *et al.*, 1995). huCdc7 was found to be localized in nuclei from the G_1 through the S phase, and no obvious relocation of huCdc7 upon progression into the S phase was detected (Figure 7G and H). More than 50% of the nuclear huCdc7 protein was extracted in buffer containing 0.5 M NaCl, but remained in the pellet after digestion with 2 mg/ml DNase I (Figure 6C). Similarly, MCM3 was extracted by salt but not by DNase I, as previously reported (data not shown; Kimura *et al.*, 1994). Therefore, it is likely that the majority of huCdc7 prepared from a random culture binds to some nuclear structures rather than to chromatin.

huCdc7 phosphorylates MCM proteins

Studies on yeast and *Xenopus* DNA replication indicated that eukaryotic replication origins may be associated with ORC and MCM protein complexes, and functions of these origin-associated proteins may be regulated in a cell cycle-specific manner (Blow, 1993; Yan *et al.*, 1993; Kimura *et al.*, 1994; Chong *et al.*, 1995; Todorov *et al.*, 1995; Coue *et al.*, 1996). Therefore, these origin-associated proteins could be the targets of phosphorylation events essential for G_1 to S transition (Bell, 1995; Carpenter *et al.*, 1996; Leatherwood *et al.*, 1996). The Cdc7–Dbf4 kinase complex in budding yeast is likely to be bound at the origins due to the origin-binding activity of Dbf4 protein (Dowell *et al.*, 1994), suggesting the possibility that ORC and MCM may be phosphorylated by Cdc7 kinase.

We examined whether MCM components can be phosphorylated by huCdc7 kinase. Extracts were prepared from

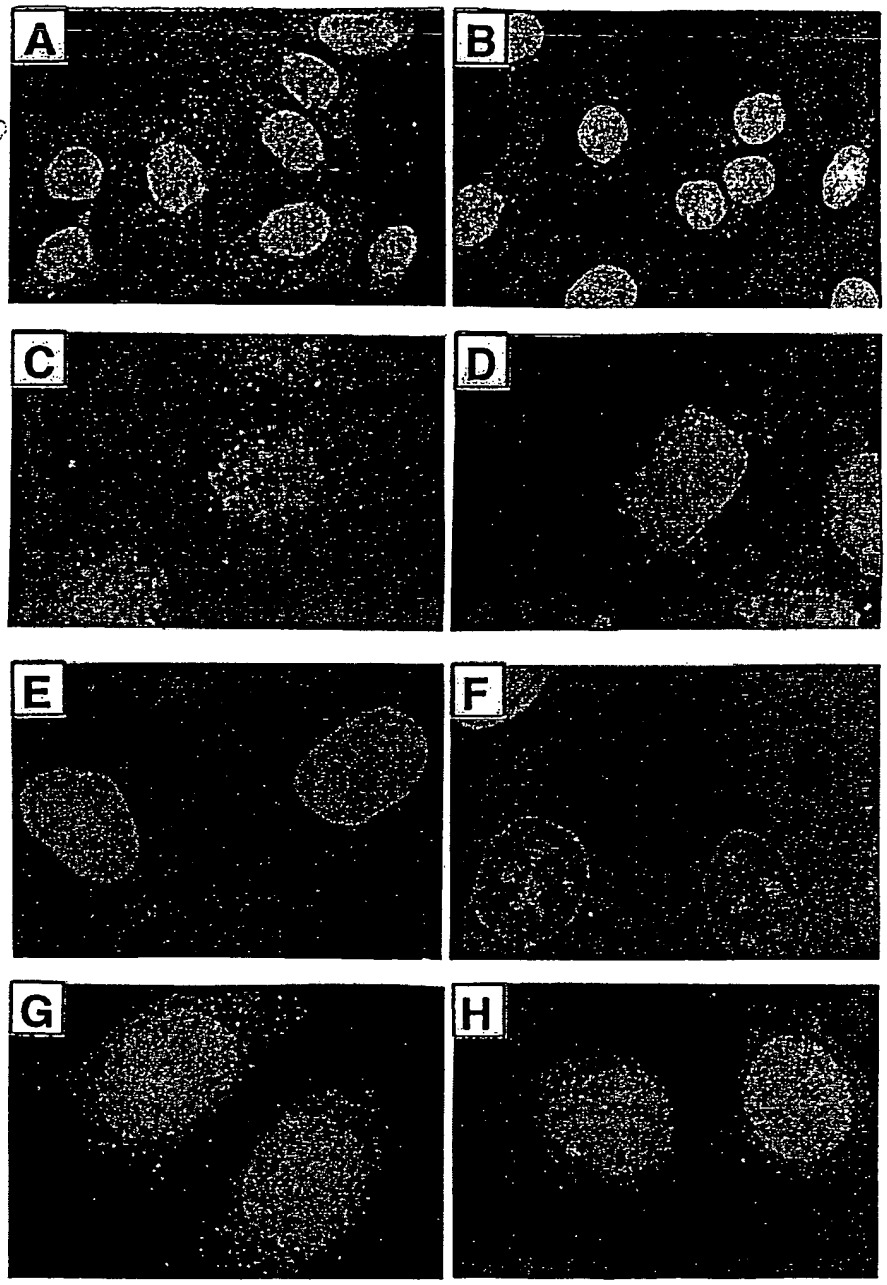


Fig. 7. Nuclear localization of huCdc7 does not alter before and after G_1/S transition. HeLa cells were synchronized at the G_1/S boundary by double thymidine block (A, C, E and G) and then released for 4 h to proceed into the S phase (B, D, F and H). (A) and (B) are superimposed images of propidium iodide staining of DNA (red) and BrdU staining (green). PCNA (C and D), MCM4 (E and F) and huCdc7 (G and H) are immunostained with specific antibodies. Magnification; 200× (A and B), 400× (C–H).

COS7 cells transfected with a vector, haemagglutinin (HA)-tagged wild-type huCDC7-expressing plasmid, or HA-tagged K90R kinase-negative huCDC7-expressing plasmid. Roughly equal amounts of transiently expressed huCdc7 protein were immunoprecipitated from the wild-type and K90R transfectants by anti-HA antibody, while huCdc7 was not detected in the immunoprecipitate from vector-transfected COS7 cells (Figure 8C). Upon incubation of the immunoprecipitates with [γ-^{32}P]ATP in the presence of purified GST fusion protein containing *Xenopus* MCM2N (amino acid residues 1–559) or human MCM3 (amino acid residues 1–808), the MCM proteins were phosphorylated efficiently by the wild-type huCdc7 immunoprecipitate, while the level of phosphorylation by K90R huCdc7 was no more higher than the vector control (Figure 8A and B). The results indicate that MCM2 and MCM3 proteins can be phosphorylated by huCdc7 *in vitro* and suggest the possibility that functions of the MCM

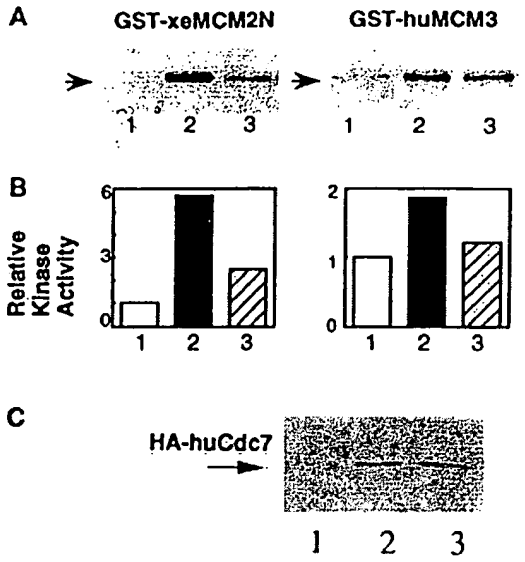


Fig. 8. *In vitro* phosphorylation of GST–MCM proteins by huCdc7 kinase. COS7 cells were transfected with 1 μg of either pKU3 vector (lane 1), pKU-HA-tagged WT huCDC7 (wild-type, lane 2) or pKU-HA-tagged K90R huCDC7 (kinase-negative form, lane 3). Two days after the transfection, cells were lysed and transiently expressed huCdc7 protein was immunoprecipitated with anti-HA antibody. (A) The immunoprecipitates were incubated with [γ-^{32}P]ATP in the presence of either GST-fused *Xenopus* MCM2N (amino acid residues 1–559) or GST-fused human MCM3 (amino acid residues 1–808). The reaction mixtures were resolved by 10% SDS–PAGE, and the gel was dried and autoradiographed. (B) ^{32}P incorporation into the MCM protein bands was quantified by Image Analyzer (Fuji Film) and the values relative to the vector control, which is taken as 1, are shown. (C) Immunoprecipitates used for the kinase assays were blotted with the anti-huCdc7 polyclonal antibody #1.

complex are regulated by phosphorylation by the Cdc7-related kinase.

Discussion

Presence of CDC7-related kinases in higher eukaryotes

Replication of eukaryotic cells is controlled precisely during the cell cycle. In spite of apparent diversity in structures of replication origins in higher eukaryotes in comparison with those of a lower eukaryote such as *S.cerevisiae*, there appears to be striking conservation in proteins required for the processes of assembly of replication machinery (Gavin *et al.*, 1995; Hamlin and Dijkwel, 1995; Donovan and Diffley, 1996). Proteins related to components of ORC, a protein complex bound specifically to the budding yeast replication origins, were discovered in *Drosophila*, *Xenopus* and human. Furthermore, they also form a multi-protein complex containing protein components similar to those of budding yeast (Gavin *et al.*, 1995; Gossen *et al.*, 1995; Carpenter *et al.*, 1996). MCM proteins, originally discovered in *S.cerevisiae*, were identified as components for the licensing factor essential for DNA replication in *Xenopus* egg extracts (Chong *et al.*, 1995; Kubota *et al.*, 1995). Proteins related to all the six MCM components have been identified in human as well (Hu *et al.*, 1993). Thus, basic mechanisms of initiation of chromosomal replication as well as its regulation may be conserved from yeasts to human. We previously reported *hsk1*$^+$, a putative homologue of CDC7 from a distantly related yeast, *S.pombe*, and suggested the possibility that the S phase initiation in eukaryotes may be regulated in a conserved manner involving Cdc7-related kinases (Masai *et al.*, 1995). Isolation of Cdc7-related kinases from human and *Xenopus*, reported in this study, further strengthens our proposal that Cdc7-related kinases are the key regulators for initiation of DNA replication conserved in eukaryotes.

The Cdc7-related kinases from higher eukaryotes share structural similarity to the yeast counterparts, exhibiting 42–44% identity in the conserved domains for serine-threonine kinases in addition to the presence of two kinase insert sequences at the conserved locations (Figure 3A). The Cdc7-related kinases were grouped into a subset distantly related to other kinases in a phylogenetic tree (Figure 3B), indicating that Cdc7-related kinases belong to a distinct kinase subfamily. Unexpectedly, huCdc7 and xeCdc7 did not carry a C-terminal acidic region, which was present and essential for the activity in the yeast genes (our unpublished data). In budding yeast Cdc7, the C-terminal tail is involved in interaction with Dbf4 protein (Patteron *et al.*, 1986). We recently discovered that efficient interaction with Dbf4 protein requires the kinase insert II and III sequences of Cdc7 (our unpublished data). Therefore, a putative 'activator' for huCdc7 may well interact with the two kinase insert sequences in the absence of the C-terminal tail. An alternative possibility is that an as yet identified variant of Cdc7-related kinases, which does contain a C-terminal region, may be present in higher eukaryotes.

Expression of huCdc7 in various tissues and in the cell cycle

mRNAs for huCdc7 are expressed in most tissues examined (Figure 5), as expected from its essential function for cell proliferation, although the highest expression was detected in testis. This may be interesting in the light of a previous report (Sclafani *et al.*, 1988) and our unpublished observations that mRNAs for budding yeast *CDC7* and fission yeast *hsk1*$^+$ are induced during the course of meiosis, suggesting that Cdc7-related kinases may play additional roles during meiosis. Characterization of budding yeast Cdc7 mutants indicated that Cdc7 is required for synaptonemal complex formation during meiosis (Sclafani *et al.*, 1988). A huCdc7 transcript was also detected in brain. The functions, if any, of huCdc7 in mature neuronal cells which do not proliferate remain to be investigated.

Among the three transcripts detected, the 3.5 kb transcript was ubiquitously present in all the tissues, whereas the 4.4 kb transcript was seen in subsets of tissues, such as testis, peripheral blood leukocytes (PBL), thymus, spleen, small intestine, brain and placenta. In various human cell lines, the former was the major transcript and the latter was expressed at a lower level (data not shown). The 2.4 kb transcript was detected only in testis, suggesting that it may be specific to this tissue. Northern analysis using the C-terminal region of huCDC7 containing only the kinase domain XI as a probe (probe B) showed that the 2.4 kb transcript did not hybridize with this DNA segment (Figure 5). In accordance with this observation,

we have obtained from a testis library a variant cDNA whose coding frame is truncated at amino acid position 519. A mouse cDNA for Cdc7-related kinase that we have isolated from a spermatocyte library contained the kinase subdomains I–VII, but its coding region was truncated in kinase insert II and continued into unrelated sequences which were identified on the mouse genomic DNA upstream of the remaining kinase domains (VII–XI), suggesting that these cDNAs are products of alternative splicing (our unpublished data). On the other hand, clone #8 contained a 91 bp insertion at the N-terminal coding region, thus resulting in frameshifting in translation. This insertion occurs at an exon–intron junction of the murine CDC7 gene (our unpublished data). At present, we do not know the functions of these apparently kinase-inactive derivatives of Cdc7-related proteins.

In lysates from human cell lines such as HeLa, TF-1 and K562, huCdc7 protein was identified as a 68 kDa protein that co-migrated with the polypeptide expressed from an expression vector carrying the 574 amino acid long huCdc7 cDNA. The level of expression of the 68 kDa huCdc7 protein did not vary significantly during the course of the mitotic cell cycle (Figure 6B). Transcription of budding yeast *CDC7* and fission yeast *hsk1*$^+$ was also previously reported to be relatively constant during the cell cycle (Yoon *et al.*, 1993; our unpublished data)

huCdc7 is a nuclear protein

Indirect immunofluorescence staining using anti-huCdc7 monoclonal antibody (4A8) showed that fluorescence was confined mostly to the nucleus during the interphase. huCdc7 was localized in nuclei before and after DNA replication is initiated, and continued to stay in nuclei during the entire interphase (Figure 7). Localization of huCdc7 in nuclei did not coincide precisely with that of PCNA, which is known to co-localize at the replication foci. Further analysis is needed to determine the precise subnuclear localization of huCdc7. huCdc7 protein could be extracted by high salt, but not by DNase I (Figure 6C), suggesting that it may be associated with nuclear structures.

Cdc7 may regulate MCM function by phosphorylation

MCM appears to be phosphorylated at various stages of the cell cycle. The newly synthesized P1 (mouse MCM3) is phosphorylated in the G_1 phase, and the level of its phosphorylation increases during the S phase (Kimura *et al.*, 1994). Phosphorylation of *Xenopus* MCM4 (Cdc21) in early S phase was also reported (Coue *et al.*, 1996). Thus, phosphorylation of MCM may activate its functions for S phase initiation. Alternatively, phosphorylation of MCM may block re-replication by dissociating from the chromatin after the initiation of S phase. Specific phosphorylation of MCM subunits by huCdc7 (Figure 8) supports the idea that this kinase regulates DNA replication. huCdc7 may regulate both activation of the S phase and 'once and only once' replication through phosphorylation of MCM subunits as well as that of as yet identified substrates. In budding yeast, it was shown recently that MCM2, 3, 4 and 6 could be phosphorylated by the Cdc7–Dbf4 kinase complex *in vitro* (A.Sugino and B.K.Tye, personal communication). Experiments are in progress to determine whether other MCM subunits are phosphorylated by huCdc7 and to locate more precisely the phosphorylation sites on MCM proteins. MCM functions could be regulated through sequential phosphorylation and dephosphorylation by multiple kinases and phosphatases. It would also be important to understand how phosphorylation of MCM by a Cdc7-related kinase is coordinated with that by other kinases, which may include Cdks and DNA-dependent protein kinase, to achieve precise regulation of MCM functions for progression of the cell cycle.

The *in vitro* kinase activity of *S.cerevisiae* Cdc7 is strictly dependent on the presence of Dbf4 protein, and Cdc7 alone expressed in insect cells is inactive. Similarly, huCdc7 alone expressed in insect cells did not show any phosphorylation activity (our unpublished data), suggesting the requirement for an activator for kinase activity of huCdc7. Although we were able to measure huCdc7-dependent phosphorylation after overexpression, we had difficulties in measuring the kinase activity of endogenous huCdc7 protein. This may reflect scarcity and/or instability of the active huCdc7 kinase complex in the cells. Identification of the putative activator for huCdc7 will help to understand the precise roles of huCdc7-dependent phosphorylation in mammalian cell cycle progression.

In summary, we report here the presence of Cdc7-related kinases in higher eukaryotes and present data implicating this kinase in regulation of mammalian chromosomal replication. Further characterization of these Cdc7-related kinases should provide important insights into molecular mechanisms of cell cycle regulation of chromosomal replication in higher eukaryotes.

Materials and methods

Cells

Mouse ES cells were cultured on a 0.1% porcine skin gelatin-coated culture dish in high glucose Dulbecco's modified Eagle's medium (DMEM) containing 0.1 mM non-essential amino acids (Gibco), 2 mM L-glutamine (Irvine Scientific), 1× nucleotide mix (3 mM each of dATP, dCTP, UTP, dGTP and 1 mM dTTP), 50 mM 2-mercaptoethanol (Sigma), the supernatant from human leukaemia-inhibiting factor (hLIF)-expressing COS7 cells and 20% fetal calf serum (FCS). COS7 and HeLa cells were cultured in DMEM containing 10% FCS. TF-1 cells were cultured in RPMI 1640 containing 2 ng/ml hGM-CSF and 10% FCS. K562 cells were cultured in RPMI 1640 containing 10% FCS.

Oligonucleotides

The oligonucleotides used for degenerate PCR amplification for cloning of CDC7-related kinases were designed on the basis of the amino acid sequences conserved between the products of *S.cerevisiae CDC7* and *S.pombe hsk1*$^+$. The combination of the oligonucleotide primers that led to the isolation of a mammalian Cdc7-related kinase was 5′-CGGAATT-CAA(AG)AT(TCA)AA(AG)GA(TC)AA(AG)AT-3′ and 5′-CGGAATT-CIGCIA(GA)ICC(AG)AA(AG)TC(ATGC)AC-3′, corresponding to the amino acid stretches from 34 to 39 and from 186 to 181, respectively, of Cdc7 protein. The other nested combination was 5′-CGGAA-TTCAA(AG)AT(TCA)GG(TCGA)GA(AG)GG(TCGA)AC-3′ and 5′-CGGGATCCIGG(TC)TT(AGT)AT(AG)TC(TCGA)C(TG)(AG)TG - 3′, corresponding to the amino acid stretches from 38 to 43 and from 166 to 161 (Patteron *et al.*, 1986).

Antibodies

Portions of the huCDC7 coding frame (amino acid residues 128–276 and 128–433), isolated as *Sau*3A fragments, were subcloned at the *Bam*HI site of pGEX-3X to generate GST fusion proteins #1 and #2, respectively, which were purified as previously described (Ikeda *et al.*, 1996a,b). Polyclonal antibodies #1 was developed in rabbit against the purified fusion protein #1 and the antibodies reacting to the GST portion

of the fusion protein were depleted by glutathione–Sepharose 4B resin to which non-fused GST protein was attached. Mouse monoclonal antibody (4A8) was developed against the GST–huCDC7 fusion protein #2. Anti-rat PCNA antibody was purchased from MBL (Nagoya, Japan) and anti-mouse Cdc21 rabbit serum was kindly provided by Dr H.Kimura (Hokkaido University, Japan). Anti-BrdU antibody containing DNase I was purchased from Amersham. Anti-HA antibody 12CA5 was purchased from Babco (CA).

GST–MCM fusion proteins

GST–human P1 (MCM3) (amino acid residues 1–808) and GST–*Xenopus* MCM2 (amino acid residues 1–559) were gifts from Dr H.Takisawa (Osaka University, Japan).

mRNA isolation and reverse transcription

Poly(A) RNA was isolated from mouse ES cells using a FAST TRACK mRNA isolation kit (Invitrogen). Reverse transcription was performed with Superscript II (Gibco) as suggested by the manufacturer. One μl of the 20 μl reaction mixture was used for subsequent PCR amplification.

PCR screening and subcloning of amplified fragments

PCR, isolation and subcloning of amplified DNA fragments were performed as described earlier (Masai *et al.*, 1995). The isolated fragments were subcloned into the T vector prepared from KS (pBluescript) vector (Marchuk *et al.*, 1991). Plasmid DNAs containing insert DNAs were recovered from white colonies of DH5α on LB plates containing ampicillin (50 μg/ml) and Xgal (40 μg/ml), and the nucleotide sequences of the inserts were determined. One of the clones carrying an insert of ~420 bp contained a coding frame that resembled a part of Cdc7 and Hsk1, which was designated pKS-420.

Screening of cDNA libraries

A human fetal liver cDNA lambda library was kindly provided by Dr Soma (Kirin Brewery Co., Japan) and a human testis cDNA lambda library was purchased from Clontech. *Xenopus* oocyte cDNA lambda libraries were kindly provided by Dr Douglas Melton and Dr Tim Hunt. A total of 10^6 plaques from each library were screened with the ^{32}P-labelled 414 bp PCR-amplified DNA fragment isolated from pKS-420. Clones #8 and #13 from the fetal liver library and #1 from the testis library carried 1.8, 2.9 and 0.8 kb cDNA inserts, respectively, and were analyzed further. Similarly, #25, #28 and #23, which carried a 1.3, 1.3 and 1.35 kb cDNA insert, respectively, were isolated from the *Xenopus* libraries. These inserts were subcloned into KS vector and the entire nucleotide sequences were determined.

Northern and genomic Southern analysis

Human multiple tissue Northern blots and zoo blot (*Eco*RI digestion) were purchased from Clontech. ^{32}P-labelled DNA probes were prepared by random priming reactions on the mixture of the 0.7 kb and 1.1 kb *Eco*RI fragments derived from the huCDC7 #8 clone containing amino acids 1–511. Hybridization was carried out in a buffer containing 6× SSPE, 10% formamide, 5× Denhardt's, 0.1% SDS and 100 μg/ ml heat-denatured sonicated salmon sperm DNA (Sigma) at 42°C for 12 h. For cross-species hybridization, the filters were washed at 42°C in 6× SSC followed by washing in 2× SSC at the same temperature. For more stringent hybridization, they were washed further in 0.2× SSC at 42°C and finally at 50°C in 0.1× SSC. The washing buffer contained 0.1% SDS.

Construction of vectors that express full-length wild-type and mutant forms of huCDC7

The 0.7 (N-terminal) and 1.1 kb (C-terminal) *Eco*RI fragments constituting the 1.8 kb insert of #8 cDNA (lacking a portion of the C-terminus coding region) were subcloned into KS vector, resulting in KS(0.7) and KS(1.1), respectively. The N-terminal coding region missing in the #13 cDNA was reconstructed by replacing the 0.25 kb *Sac*I–*Bcl*I fragment of #13 with that (0.6 kb) of KS(0.7), generating KS-huCDC7-#6 which contained the entire coding frame of huCDC7 with a 91 nucleotide insertion at amino acid position 39 as well as the 3'-untranslated region. The spliced form containing the entire coding region in-frame was constructed by replacing the *Age*I–*Kas*I fragment of KS-huCDC7-#6 (containing the 91 nucleotide insert) with that of testis-derived #1 cDNA (without the insert), resulting in KS-full huCDC7. The N-terminal *Not*I–*Age*I fragment of the KS-full huCDC7 was replaced by the fragment generated by *Not*I–*Age*I digestion of a PCR-amplified DNA, resulting in the HA-tagged 'short' huCdc7 coding frame (starting from amino acid position 13; KS-short huCDC7). This PCR was conducted by using an oligonucleotide containing *Not*I–*Nde*I sites followed by the sequences encoding the 10 amino acid HA peptide (MYPYDVPDYA) and the huCdc7 coding region (amino acids 13–17) in combination with an internal primer 5'-TTT GTC CTC AAT CTT AAT CTT-3' present downstream of the *Age*I site. The 2.3 kb *Not*I–*Xba*I fragment of KS-short huCDC7 containing the HA-tagged short form was cloned into pKU3 (a gift from Dr Muto of our laboratory), a neomycin-resistant derivative of pME18S, resulting in pKU3-HA-short huCDC7 which expressed HA-tagged 562 amino acid huCDC7 under the SRα promoter. For construction of a plasmid expressing the full-length huCDC7, the 180 bp *Eco*RI–*Age*I fragment of KS-full huCDC7 containing the N-terminal huCDC7 coding frame (from amino acid position 1 to 22) was further digested by *Alu*I (at 31 nucleotides upstream of the first ATG), and a *Not*I linker was attached to this *Alu*I site. This *Not*I–*Age*I fragment replaced the same fragment of pKU3-HA-short huCDC7, resulting in pKU3-long huCDC7 containing the 574 amino acid full-length huCDC7 coding frame. To generate pKU3-short huCDC7, the same *Not*I–*Age*I fragment of pKU3-HA-short huCDC7 was replaced by an oligonucleotide containing a *Not*I site followed by the sequence encoding 10 amino acids from position 13 to 22 of huCDC7.

A mutant form of huCDC7 in which the lysine at position 90 was replaced by arginine was constructed by the PCR method. Oligonucleotides 5'-AAAATTGCCTTAAGACACTTGATTCCAACA-3' and 5'-TCAAGTGTCTTAAGGCAATTTTCTCTTCAG-3' were used to create the mutation.

Transfection of plasmid DNA into mammalian cells

Plasmid DNAs were introduced into COS7 cells by electroporation as previously described (Kitamura *et al.*, 1991).

Cell synchronization and preparation of cell lysates

HeLa cells were synchronized at the G_1/S boundary by double thymidine block as described (O'Connor and Jackman, 1995). Factor-depleted TF-1 cells were prepared by deprivation of hGM-CSF for 16 h and the G_1 phase TF-1 cells were obtained by mimosine treatment as described (O'Connor and Jackman, 1995). Cells were synchronized in metaphase by nocodazole treatment (O'Connor and Jackman, 1995). Synchronization of the cell cycle was monitored by flow cytometry (Giunta and Pucillo, 1995). Cells were washed with ice-cold phosphate-buffered saline (PBS) and were resuspended in IP buffer [50 mM HEPES/KOH (pH 7.6), 150 mM NaCl, 2.5 mM EGTA, 1 mM EDTA, 1 mM dithiothreitol (DTT), 0.1% Tween-20, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride, 1 μg/ml aprotinin, 1 μg/ml leupeptin, 100 μM β-glycerophosphate, 10 mM NaF and 1 mM Na_3VO_4] at 2×10^7 cells/ml. After sonication of the cell suspension, lysates were cleared by centrifugation. Preparation of nuclear extracts and salt/DNase I extraction were conducted as described by Kimura (1994).

Immunoprecipitaion, immunoblot and in vitro kinase assays

Anti-huCDC7 antibody #1 or anti-HA antibody (12CA5) was added to the cell lysate at a final concentration of 1 μg/ml and incubated for 2 h on ice. Immunoprecipitates were collected on protein A–agarose beads, washed three times with IP buffer. For immunoblot analysis, bound proteins were extracted by boiling with 20–40 μl of 2× Laemmli's sample buffer. Proteins were separated on an 8% SDS–PAGE and were blotted onto an Immobilon-P membrane (Millipore). The membranes were probed with anti-huCDC7 followed by incubation with horseradish peroxidase-conjugated anti-rabbit IgG (Amersham). The immunoreactive proteins were detected by chemiluminiscence (ECL, Amersham, UK). For *in vitro* kinase assays, immunoprecipitates were washed further by pre-kinase buffer [40 mM HEPES/KOH (pH 8.0) and 40 mM potassium glutamate] and incubated in 24 μl of kinase reaction buffer [40 mM HEPES/KOH (pH 8.0), 40 mM potassium glutamate, 1 mM EGTA, 8 mM magnesium acetate, 2 mM DTT, 0.5 mM EDTA, 0.1 mM ATP and 2 μCi of [γ-^{32}P]ATP] in the presence or absence of 1 μg of GST-fused MCM protein at 30°C for 20 min. The reaction was stopped by addition of 6 μl of 5× Laemmli's sample buffer. Samples were heated at 95°C for 5 min and were separated on a 10% SDS–PAGE. The gels were dried and phosphorylated proteins were detected by autoradiography.

Indirect immunofluorescence

Cells grown on coverslips were washed twice with PBS containing 1 mM $CaCl_2$ and 1.5 mM $MgCl_2$. Cells were fixed for 5 min with 3.7% formaldehyde and permeabilized with 0.2% Triton X-100 in PBS for 5 min followed by incubation in blocking solution (0.2% gelatin, 5% FCS and 0.1% Tween-20 in PBS) when extraction of soluble proteins was not necessary. For BrdU staining, cells were incubated with 20 μM BrdU for 20 min prior to the staining. In case soluble proteins needed

to be extracted, cells were washed twice with PBS and once with CSK buffer [100 mM NaCl, 300 mM sucrose, 10 mM PIPES (pH 6.8) and 3 mM $MgCl_2$] (Fey *et al.*, 1984), permealized by 0.5% Triton X-100 in CSK buffer and fixed with 3.7% formaldehyde at room temperature for 5 min. The coverslips were incubated in blocking solution for 10 min and for a further 1 h after addition of the first antibody (5 μg/ml) and washed three times with the same solution. They were incubated further with a second antibody [rhodamine-conjugated goat anti-rabbit antibody or rhodamine-conjugated goat anti-mouse antibody (Immunotech, France) or fluorescein isothiocyanate-conjugated rabbit anti-mouse antibody (Zymed)] which had been diluted 1:250 in blocking solution. Finally, the coverslips were washed three times with PBS and then visualized under the immunofluorescence microscopy (Nikon Optiphot-2).

Accession numbers

The DDBJ/EMBL/GenBank accession numbers for huCDC7 and XeCDC7 reported here are AB003698 and AB003699, respectively.

Acknowledgements

We are grateful to Dr Haruhiko Takisawa for the generous gift of GST–human MCM3 and GST–*Xenopus* MCM2 proteins; Drs Julian Blow and H.M.Mahbubani for conducting a part of the screening for *Xenopus* CDC7 cDNAs; Hiroyuki Kumagai for determining the sequence of the 3′-untranslated region of huCDC7 cDNA; Masayuki Yamada for helping in sequencing xeCDC7 cDNA; Takahisa Hachiya and Takashi Moritsu of MBL for development of rabbit and mouse antibodies; Drs H.Doi and Y.Mizusawa for a helpful suggestion on the use of the Fujitsu SINCA program; and Dr Masaaki Muramatsu, Dr Tsuyoshi Miyake, Dr Akihira Ohtoshi and Kim Jung Min for helpful suggestions and discussions. This work was supported partly by a Grant-in-Aid for Scientific Research on Priority Research from the Ministry of Education, Science, Sports and Culture of Japan to H.M. and by a Grant-in-Aid for Encouragement of Young Scientists to N.S. from the same Ministry.

References

Bell,S.P. (1995) Eukaryotic replicators and associated protein complexes. *Curr. Opin. Genet. Dev.*, 5, 162–167.

Bell,S.P. and Stillman,B. (1992) Nucleotide dependent recognition of chromosomal origins of DNA replication by a multi-protein complex. *Nature*, 357, 128–134.

Bell,S.P., Kobayashi,R. and Stillman,B. (1993) Yeast origin recognition complex (ORC) in transcriptional silencing and DNA replication in *S.cerevisiae*. *Science*, 262, 1844–1849.

Blow,J.J. (1993) Preventing re-replication of DNA in a single cell cycle: evidence for a replication licensing factor. *J. Cell Biol.* 122, 993–1002.

Blow,J.J. amd Laskey,R.A. (1988) A role for the nuclear envelope in controlling DNA replication within the cell cycle. *Nature*, 332, 546–548.

Campbell,J.L. and Newlon,C.S. (1991) Chromosomal DNA replication. In Broach,J.R., Pringe,J.R. and Jones,E.W. (eds), *The Molecular and Cellular Biology of the Yeast Saccharomyces*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 41–146.

Carpenter,PB., Mueller,P.R. and Dunphy,W.G. (1996) Role for a *Xenopus* Orc2-related protein in controlling DNA replication. *Nature*, 379, 357–360.

Chong,J.P., Mahbubani,H.M., Khoo,C.Y. and Blow,J.J. (1995) Purification of an MCM-containing complex as a component of the DNA replication licensing system. *Nature*, 375, 418–421.

Cocker,J.H., Piatti,S., Santocanale,C., Nasmyth,K. and Diffeley,J.F.X. (1996) An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. *Nature*, 379, 180–182.

Coue,M., Kearsey,S.E. and Mechali,M. (1996) Chromatin binding, nuclear localization and phosphorylation of *Xenopus* cdc21 are cell-cycle dependent and associated with the control of initiation of DNA replication. *EMBO J.*, 15, 1085–1097.

Coverley,D. and Laskey,R.A. (1994) Regulation of eukaryotic DNA replication. *Annu. Rev. Biochem.*, 63, 745–776.

Diffley,J.F.X. and Cocker,J.H. (1992) Protein–DNA interactions at a yeast replication origin. *Nature*, 357, 169–172.

Diffley,J.F.X., Cocked,J.H., Dowell,S.J. and Rowley,A. (1994) Two steps in the assembly of complexes at yeast replication origins *in vivo*. *Cell*, 78, 303–316.

Donovan,S. and Diffley,J.F.X. (1996) Replication origins in eukaryotes. *Curr. Opin. Genet. Dev.*, 6, 203–207.

Dowell,S.J., Romanowski,P. and Diffley,J.F.X. (1994) Interaction of Dbf4, the Cdc7 protein kinase regulatory subunit, with yeast replication origins *in vivo*. *Science*, 265, 1243–1246.

Ehrenhofer-Murray,A.E., Gossen,M., Pak,D.T., Botchan,M.R. and Rine,J. (1995) Separation of origin recognition complex functions by cross-species complementation. *Science*, 270, 1671–1674.

Fey,E.G., Wan,K.M. and Penman,S. (1984) Epithelial cytoskeletal framework and nuclear matrix-intermediate filament scaffold: three-dimensional organization and protein composition. *J. Cell Biol.*, 98, 1973–1984.

Fox,C.A., Loo,S., Dillin,A. and Rine,J. (1995) The origin recognition complex has essential functions in transcriptional silencing and chromosomal replication. *Genes Dev.*, 9, 911–924.

Gavin,K.A., Hidaka,M. and Stillman,B. (1995) Conserved initiator proteins in eukaryotes. *Science*, 270, 1667–1671.

Giunta,M. and Pucillo,C. (1995) Techniques for monitoring cell cycle phase. In Pagano,M. (ed.), *Cell Cycle—Material and Methods*. Springer Laboratory Manual, Cambridge, MA, pp. 63–74.

Gossen,M., Pak,D.T., Hansen,S.K., Acharya,J.K. and Botchan,M.R. (1995) A *Drosophila* homolog of the yeast origin recognition complex. *Science*, 270, 1674–1677.

Hamlin,J.L. and Dijkwel,A. (1995) On the nature of replication origins in higher eukaryotes. *Curr. Opin. Genet. Dev.*, 5, 153–161.

Hanks,S.K., Quinn,A.M. and Hunter,T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science*, 241, 42–52.

Hartwell,L.H. (1974) *Saccharomyces cerevisiae* cell cycle. *Bacteriol Rev.*, 38, 164–98.

Hu,B., Burkhart,R., Schulte,D., Musahl,C. and Knippers,R. (1993) The P1 family: a new class of nuclear mammlian proteins related to the yeast Mcm replication proteins. *Nucleic Acids Res.*, 21, 5289–5293.

Huberman,J.A. (1995) Prokaryotic and eukaryotic replicons. *Cell*, 82, 535–542.

Ikeda,M., Arai,K.-i. and Masai,H. (1996a) CTBP1/RBP1, a *Saccharomyces cerevisiae* protein which binds to T-rich single stranded DNA containing the 11-bp core sequence of autonomously replicating sequence, is a poly (deoxypyrimidine)-binding protein. *Eur. J. Biochem.*, 238, 38–47.

Ikeda,M., Arai,K.-i. and Masai,H. (1996b) A fusion protein library: an improved method for rapid screening and characterization of DNA binding or interacting proteins. *Gene*, 181 167–171.

Jackson,A.L., Pahl,P.M., Harrison,K., Rosamond,J. and Sclafani,R.A. (1993) Cell cycle regulation of the yeast Cdc7 protein kinase by association with the Dbf4 protein. *Mol. Cell. Biol.*, 13, 2899–2908.

Kearsey,S.E., Labib,K. and Maiorano,D. (1996) Cell cycle control of eukaryotic DNA replication. *Curr. Opin. Genet. Dev.*, 6, 208–214.

Kimura,H., Nozaki,N. and Sugimoto,K. (1994) DNA polymerase α associated protein P1, a murine homologue of yeast MCM3, changes its intranuclear distribution during the DNA synthesis period. *EMBO J.*, 13, 4311–4320.

Kimura,H., Takizawa,N., Nozaki,N and Sugimoto,K. (1995) Molecular cloning of cDNA encoding mouse Cdc21 and CDC46 homologs and characterization of the products: physical interaction between P1 (MCM3) and CDC46 proteins. *Nucleic Acids Res.*, 23, 2097–2104.

Kitada,K., Johnston,L.H., Sugino,T. and Sugino,A. (1992) Temperature-sensitive cdc7 mutations of *Saccharomyces cerevisiae* are suppressed by the DBF4 gene, which is required for the G1/S cell cycle transition. *Genetics*, 131, 21–29.

Kitamura,T., Sato,N., Arai,K. and Miyajima,A. (1991) Expression cloning of the human IL-3 receptor cDNA reveals a shared beta subunit for the human IL-3 and GM-CSF receptors. *Cell*, 66, 1165–1174.

Kubota,Y., Mimura,S., Nishimoto,S.I., Takisawa,H. and Nojima,H. (1995) Identification of the yeast Mcm3-related protein as a component of *Xenopus* DNA replication licensing factor. *Cell*, 81, 601–609.

Leatherwood,J., Lopez-Girona,A. and Russell,P. (1996) Interaction of Cdc2 and Cdc18 with a fission yeast ORC2-like protein. *Nature*, 379, 360–363.

Liang,C., Weinreich,M. and Stillman,B. (1995) ORC and Cdc6p interact and determine the frequency of initiation of DNA replication in the genome. *Cell*, 81, 667–676.

Loo,S., Fox,C.A., Rine,J., Kobayashi,R. Stillman,B. and Bell,S. (1995) The origin recognition complex in silencing, cell-cycle progression, and DNA-replication. *Mol. Biol. Cell*, 6, 741–756.

Marahrens,Y. and Stillman,B. (1992) A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science*, **255**, 817–823.

Marchuk,D., Drumm,M., Saulino,A. and Collins,F. (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.*, **19**, 1154.

Masai,H., Miyake,T. and Arai,K.-i. (1995) *hskI*$^{+}$, a *Schizosacchramoces pombe* gene related to *Saccharamomyces cerevisiae CDC7*, is required for chromosomal replication. *EMBO J.*, **14**, 3094–3104.

O'Connor,P. and Jackman,J. (1995) Synchronization in mammalian cells. In Pagano,M. (ed.), *Cell Cycle—Material and Methods*. Springer Laboratory Manual, Cambridge, MA, pp. 63–74.

Patteron,M.N., Sclafani,R.A., Fangman,W.L. and Rosamund,J. (1986) Molecular characterization of cell cycle gene CDC7 from *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **6**, 1590–1598.

Rao,H. and Stillman,B. (1995) The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. *Proc. Natl Acad. Sci. USA*, **92**, 2224–2228.

Rao,H., Marahrens,Y. and Stillman,B. (1994) Functional conservation of multiple elements in yeast chromosomal replicators. *Mol. Cell. Biol.*, **14**, 7643–7651.

Rowley,A., Dowell,S.J. and Diffely,J.F.X. (1994) Recent developments in the initiation of chromosomal DNA replication: a complex picture emerges. *Biochim. Biophys. Acta*, **1217**, 239.

Rowley,A., Cocker,J.H., Harwood,J. and Diffley,J.F.X. (1995) Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. *EMBO J.*, **14**, 2631–2641.

Sclafani,R.A. and Jackson,A.L. (1994) Cdc7 protein kinase for DNA metabolism comes of age. *Mol. Microbiol.*, **11**, 805–810.

Sclafani,R.A., Patterson,M., Rosamond,J. and Fangman,W.L. (1988) Differential regulation of the yeast CDC7 gene during mitosis and meiosis. *Mol. Cell. Biol.*, **8**, 293–300.

Theis,J.F. and Newlon,C.S. (1994) Domain B of ARS307 contains two functional elements and contributes to chromosomal replication origin function. *Mol. Cell. Biol.*, **14**, 7652–7659.

Thommes,P., Fett,R., Schray,B., Burkhart,R., Barnes,M., Kennedy,C., Brown,N.C. and Knippers,R. (1992) Properties of the nuclear P1 protein, a mammalian homologue of the yeast Mcm3 replication protein. *Nucleic Acids Res.*, **20**, 1069–1074.

Todorov,I., Attaran,A. and Kearsey,S. (1995) BM28, a human member of the MCM2–3–5 family, is displaced from chromatin during DNA replication. *J. Cell Biol.*, **129**, 1433–1446.

Treisman,J.E., Follette,P.J., O'Farrell,P.H. and Rubin,G.M. (1995) Cell proliferation and DNA replication defects in a *Drosophila* MCM2 mutant. *Genes Dev.*, **9**, 1709–1715.

Walker,S.S., Malik,A.K. and Eisenberg,S. (1991) Analysis of the interactions of functional domains of a nuclear origin of replication from *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **19**, 6255.

Yan,H., Merchant,M. and Tye,B.K. (1993) Cell cycle-regulated nuclear localization of MCM2 and MCM3, which are required for the initiation of DNA synthesis at chromosomal replication origins in yeast. *Genes Dev.*, **7**, 2149–2160.

Yoon,H.J., Loo,S. and Campbell,J.L. (1993) Regulation of *Saccharomyces cerevisiae* CDC7 function during the cell cycle. *Mol. Biol. Cell*, **4**, 195–208.

Received on November 11, 1996; revised on April 1, 1997

In re: Meyers
Appl. No. 09/464,039
Filed December 15, 1999

APPENDIX F

0026-895X/01/5903-427–433$3.00
MOLECULAR PHARMACOLOGY

Mol Pharmacol 59:427–433, 2001

Vol. 59, No. 3
617/892595
Printed in U.S.A.

ACCELERATED COMMUNICATION

Discovery of a Novel Member of the Histamine Receptor Family

TUAN NGUYEN, DAVID A. SHAPIRO, SUSAN R. GEORGE, VINCENT SETOLA, DENNIS K. LEE, REGINA CHENG, LAURA RAUSER, SAMUEL P. LEE, KEVIN R. LYNCH, BRYAN L. ROTH, and BRIAN F. O'DOWD

The Centre for Addiction and Mental Health, Toronto, Ontario, Canada (T.N., S.R.G., R.C., B.F.O.); Departments of Pharmacology (S.R.G., D.K.L., S.P.L., B.F.O.) and Medicine (S.R.G.), University of Toronto, Toronto, Ontario, Canada; Department of Biochemistry (D.A.S., V.S., B.L.R.) and National Institute of Mental Health Psychoactive Drug Screening Program (L.R., B.L.R.), Case Western Reserve University Medical School, Cleveland, Ohio; and Department of Pharmacology, University of Virginia Health Sciences Center, Charlottesville, Virginia (K.R.L).

Received October 4, 2000; accepted January 5, 2001

This paper is available online at http://molpharm.aspetjournals.org

ABSTRACT

We report the discovery, tissue distribution and pharmacological characterization of a novel receptor, which we have named H4. Like the three histamine receptors reported previously (H1, H2, and H3), the H4 receptor is a G protein-coupled receptor and is most closely related to the H3 receptor, sharing 58% identity in the transmembrane regions. The gene encoding the H4 receptor was discovered initially in a search of the GenBank databases as sequence fragments retrieved in a partially sequenced human genomic contig mapped to chromosome 18. These sequences were used to retrieve a partial cDNA clone and, in combination with genomic fragments, were used to determine the full-length open reading frame of 390 amino acids. Northern analysis revealed a 3.0-kb transcript in rat testis and intestine. Radioligand binding studies indicated that the H4 receptor has a unique pharmacology and binds [^{3}H]histamine (K_d = 44 nM) and [^{3}H]pyrilamine (K_d = 32 nM) and several psychoactive compounds (amitriptyline, chlorpromazine, cyproheptadine, mianserin) with moderate affinity (K_i range of 33–750 nM). Additionally, histamine induced a rapid internalization of HA-tagged H4 receptors in transfected human embryonic kidney 293 cells.

Histamine is a monoamine neurotransmitter thus far known to activate three G protein-coupled receptors (GPCRs), the H1, H2, and H3 receptors (Hill et al., 1997). Although molecular cloning has made possible the identification, isolation, and characterization of the majority of known GPCRs, the histamine receptor subtypes have proven more difficult to identify. The first two histamine receptor genes cloned were H1 (Yamashita et al., 1991) and H2 (Gantz et al., 1991). The identification of the H3 receptor came nearly a decade later (Lovenberg et al., 1999). Collectively, the H1, H2, and H3 receptors share less than 35% identity with one another and each has greater sequence identities with other aminergic receptors. Thus, the histamine receptor gene family is significantly divergent and may have evolved from different ancestral genes (Leurs et al., 2000). In addition, there is evidence that multiple subtypes of the H3 receptor may exist. Pharmacological studies performed on membranes extracted from rat brain tissue revealed two classes of H3 binding sites (West et al., 1990; Leurs et al., 1996). However, efforts to clone a second H3 receptor subtype in the brain have thus far been unsuccessful.

Through molecular cloning techniques, we have identified numerous novel GPCRs, including many subtypes not suspected to exist on the basis of pharmacology. Over the past decade, our cloning efforts have identified a number of GPCR genes, including genes encoding such receptors as the cysteinyl leukotriene CysLT2 (Heise et al., 2000), galanin GalR2 and GalR3 (Kolakowski et al., 1998), thyrotropin-releasing hormone TRH-R2 (O'Dowd et al., 2000), uridine nucleotide receptor UNR (Nguyen et al., 1995), as well as a large cohort of orphan GPCRs for which the endogenous ligands remain to be elucidated (Marchese et al., 1999; Lee et al., 2001). In a

This research was funded by the Canadian Institutes of Health Research (B.F.O. and S.R.G.), the National Institute on Drug Abuse (B.F.O. and S.R.G.), and in part by K02-MH01366 and N01–80005 to B.L.R.

T.N. and D.A.S. contributed equally to this work.

ABBREVIATIONS: GPCR(s), G protein-coupled receptor(s); HTGS, high-throughput genomic sequence; TM, transmembrane; NCBI, National Center for Biotechnology Information; PCR, polymerase chain reaction; HA, hemagglutinin; HEK, human embryonic kidney cells; kb, kilobase pair.

scan of the GenBank high-throughput genomic sequence (HTGS) database, we identified a small DNA fragment that had greatest homology to the known histamine receptor genes. Here we report the discovery, tissue distribution, and pharmacological characterization of human DNA encoding a receptor, H4. H4 receptor mRNA had discrete and limited expression in rat testis and intestine. Although the H4 receptor was most closely related to the H3 receptor [58% identity in the transmembrane (TM) regions], it seemed to possess a unique pharmacology, with highest affinity for psychoactive drugs (amitriptyline, chlorpromazine) with a tricyclic structure.

Materials and Methods

Database Searching. A customized search was used to query the GenBank HTGS database of high-throughput genomic sequences maintained by the National Center for Biotechnology Information (NCBI) with the sequences of various GPCRs, using the FAST_PAN program (Retief et al., 1999).

Screening and Construction of the H4 Receptor Gene. Partial sequences encoding the H4 receptor were used to design primers to PCR-amplify H4-encoding fragments from human genomic DNA. A fragment encoding the start methionine to TM2 was amplified using primers P1 (5′-ATGCCAGATACTAATAGCACAATC-3′) and P2 (5′-CACAAAGAAGTCAGAGATGG-3′) and another fragment encoding from TM5 to TM6 using primers P3 (5′-TGGTACATCCTTGCCATC-3′) and P4 (5′-TATGGAGCCCAGCAAACAG-3′). PCR products were extracted with phenol/chloroform, precipitated with ethanol, and electrophoresed on a low-melting agarose gel. DNA in the expected size range were excised from the gel, ligated into the *Eco*RV site of the pBluescript vector (Stratagene, La Jolla, CA), and then the sequence was determined. H4 receptor-encoding fragments were used to screen human and rat genomic libraries (CLONTECH, Palo Alto, CA) and a human testis cDNA library (CLONTECH), as described previously (Marchese et al., 1994). Isolated phages from the human and rat genomic libraries were purified and subcloned, and the sequence was determined as described previously (Marchese et al., 1994). Isolated phages from the human cDNA library were subjected to PCR amplification using primers specific for regions flanking the insert of the λgt10 library vector, subcloned, and sequenced.

A DNA fragment encoding the full-length human H4 receptor was amplified by PCR in three stages using isolated human genomic and cDNA library phage as templates. In stage 1, three overlapping fragments (A, B, and C), together encoding the full-length H4 receptor, were PCR-amplified as follows. Fragment A was obtained using primers P5 (5′-CATCATTTGATGTGATGCCA-3′) and P6 (5′-CAAAGGAATGGAGATCACACCCACAAAGAAGTCAGA-3′) from an isolated genomic library phage obtained with the Met-TM2 encoding DNA probe. Fragment B was obtained using primers P7 (5′-GTGATCTCCATTCCTTTG-3′) and P8 (5′-TCCAATAAATATTCAT-3′) from an isolated testis cDNA library phage. Fragment C was obtained using primers P9 (5′-TGGTACATCCTTGCCATC-3′) and P10 (5′-GAGGTGAGAAAATTGTC-3′) from an isolated genomic library phage obtained with the TM5–6 encoding DNA probe. In stage 2, fragments B and C were joined by PCR using primers P7 and P10. In stage 3, fragments A and B/C were joined together using primer P5 and P10 to obtain a fragment with a length of ~1200 base pairs encoding the full-length H4 receptor. DNA encoding a hemagglutinin (HA) epitope tag (YPYDVPDYA) was inserted after the start methionine codon using PCR mutagenesis. This DNA was ligated into the *Eco*RV site of the mammalian expression vector pcDNA3 (Invitrogen, Carlsbad, CA) and its sequence was determined.

Northern Expression Analyses. mRNA from various human and rat tissues were extracted as described previously (Marchese et al., 1994). Briefly, poly(A)$^+$ RNA was isolated using oligo(dT) cellulose spin columns (Pharmacia, Uppsala, Sweden), denatured and size fractionated on a 1% formaldehyde agarose gel, transferred onto nylon membrane, and immobilized by UV irradiation. The blots were hybridized with ^{32}P-labeled DNA fragments encoding the human and rat H4 receptor, washed with 2× standard saline/phosphate/EDTA and 0.1% SDS at 50°C for 20 min and with 0.1× standard saline/phosphate/EDTA and 0.1% SDS at 50°C for 2 h and exposed to X-ray film at −70°C in the presence of an intensifying screen.

Radioligand Binding Studies. For these studies, the HA-tagged H4 receptor was transiently transfected into human embryonic kidney 293 cells into 100-mm plates using FUGENE-6 (Roche

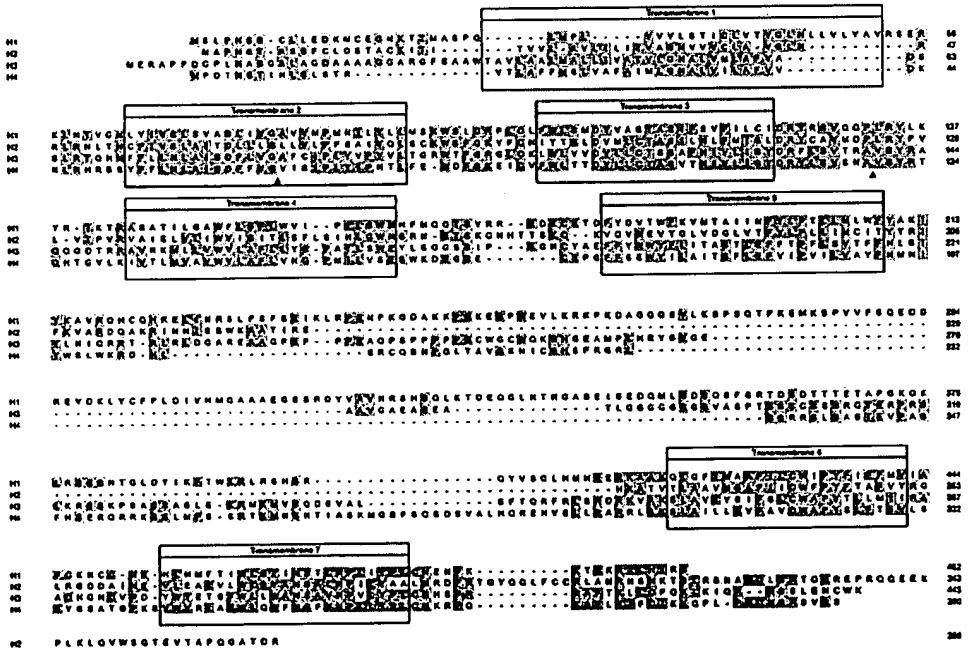


Fig. 1. Alignment of four human histamine receptors H1, H2, H3 and H4. Residues identical among the receptors are shaded. Numeric amino acid positions are indicated on the right. The presence of introns interrupting the H4 sequence are shown as "▲". H4 sequence data has been deposited in GenBank (accession number AY008280).

Molecular Biochemicals, Rotkreuz, Switzerland) at a 6:1 ratio of FUGENE/DNA as detailed previously (Shapiro et al., 2000). At 72 h after transfection, cells were harvested and membranes prepared as described previously (Shapiro et al., 2000); membranes were stored as pellets at −80°C until use. Radioligand binding assays were performed in a total volume of 500 μl in 50 mM Tris-Cl, 0.5 mM EDTA, pH 7.4, with 15 nM [^{3}H]pyrilamine in the 96-well format. After a 1-h incubation at room temperature, membranes were harvested by rapid filtration with a Brandel Harvester followed by two quick washes of ice-cold binding buffer. After drying, filters were placed into scintillation fluid and quantified by liquid scintillation spectrometry. Data were analyzed using GraphPad Prism (GraphPad Software, San Diego, CA) and data presented represents the mean ± S.E.M. of computer-derived estimates of at least three separate experiments, each done in duplicate. For inhibition studies, 12 concentrations of unlabeled ligand spanning 6 log units of test drug were used; for saturation studies, six concentrations of labeled ligand spanning 3 log units were used.

Immunoblot Analysis. The HA-tagged H4 receptor was transiently transfected into COS-7 monkey kidney cells (American Type Culture Collection, Manassas, VA) and membranes prepared from these cells as described previously (Lee et al., 2000). In brief, tissues were solubilized in sample buffer consisting of 50 mM Tris-HCl, pH 6.5, 1% SDS, 10% glycerol, 0.003% bromphenol blue, and 10% 2-mercaptoethanol. The samples were subjected to polyacrylamide gel electrophoresis with 12% acrylamide gels and electroblotted onto nitrocellulose as described previously (Ng et al., 1996). HA-tagged H4 immunoreactivity was revealed with the 3F10 rat monoclonal antibody (Roche, Laval, Quebec).

Internalization Studies. For these studies, HA-tagged H4 receptors were transfected into HEK-293 cells as described above into 100-mm plates. At 24 h after transfection, cells were split into 24-well plates containing poly-lysine-coated glass cover slips using Dulbecco's modified Eagle's medium containing 10% dialyzed fetal calf serum. The medium was replaced 24 h later with serum-free Dulbecco's modified Eagle's medium. The next day, cells were exposed to 100 μM histamine for 0, 2, 5, or 15 min and then fixed with freshly prepared 4% paraformaldehyde in PBS for 30 min at room temperature. Cells were then permeabilized on ice (0.2% Triton X-100 in ice-cold PBS) for 20 min and then incubated with blocking buffer (5% nonfat dry milk, 2% bovine serum albumin in PBS) for 1 h and then incubated with a 1:2000 dilution of monoclonal anti-HA antibody in blocking buffer overnight at 4°C. The next day, after extensive washing in room temperature PBS, cells were incubated with secondary antibody (BODIPY-labeled goat anti-mouse; 1:250 dilution in blocking buffer) for 1 h, washed extensively with PBS, and prepared for confocal microscopy as detailed previously (Berry et al., 1996; Kristiansen et al., 2000). Internalization was quantified as described previously (Berry et al., 1996; Willins et al., 1999).

Results and Discussion

As part of our ongoing search of novel genes encoding GPCRs, we queried the GenBank sequence databases maintained by NCBI with known GPCR sequences. A search with the histamine H3 receptor sequence retrieved partial sequences of a novel GPCR-encoding gene in an unfinished sequence of a human contig mapped to chromosome 18 (GenBank accession number AC007922). The retrieved sequence was obtained in three separate fragments, including one fragment that seemed to encode the receptor from the start methionine to TM2, another fragment encoding TM3, and a third fragment encoding TM5 through TM7 of a novel GPCR-like gene. These partial sequences were used to design primers for PCR amplification of human genomic DNA. Two frag-

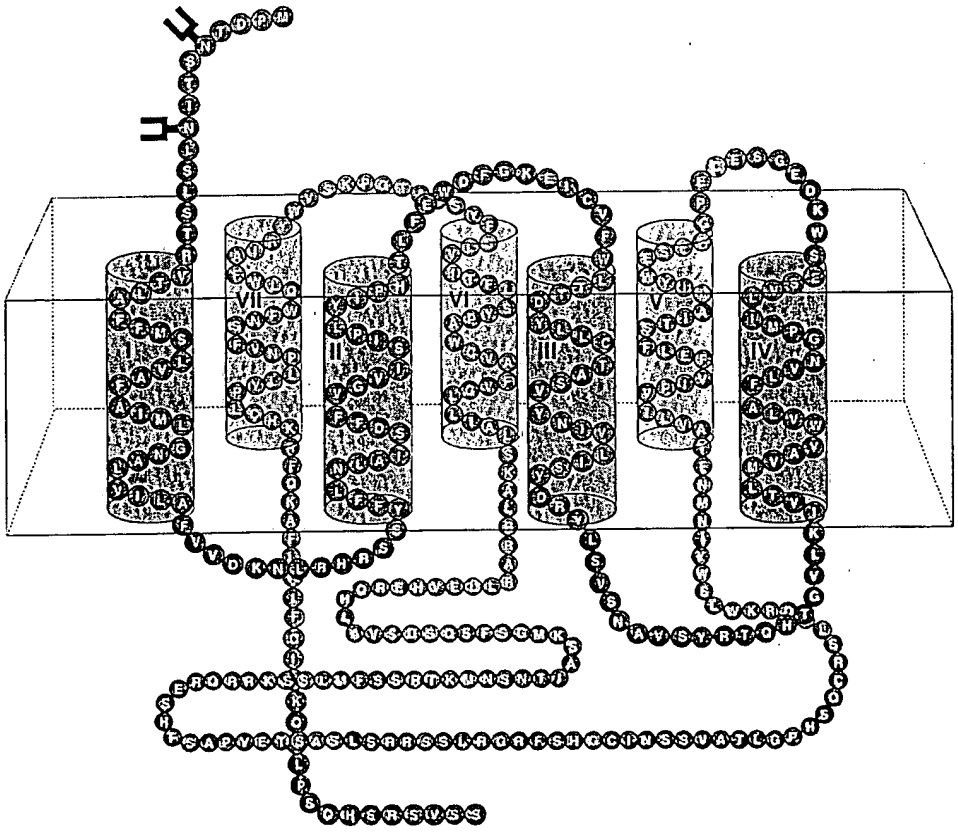


Fig. 2. Representative schematic of the human H4 receptor embedded in a cell membrane (box). Transmembrane regions are numbered and depicted with a top (extracellular)/bottom (intracellular) orientation. *N*-linked glycosylation sites are indicated in blue. Amino acids in red represent residues shared with the histamine receptor H3.

ments of DNA, encoding the start methionine to TM2 and TM5 to TM6, respectively, were obtained and used as probes to screen a human genomic library. The Met-TM2 probe retrieved two phages encoding the 5′ end of the novel gene from the start methionine to TM2, whereas the TM5-TM6 probe retrieved four phages encoding the 3′ end of the gene from TM5 to the stop codon.

Using the TM5-TM6 probe, we screened several tissues by Northern analyses, which revealed a 3.0-kb signal in rat testis. A human testis cDNA library was screened with both probes, which isolated two identical phages encoding the sequence from TM2 to TM6 of the gene. The sequences were identical in regions of overlap, confirming them to be partial sequences of the same gene. A comparison of the cDNA fragment to the GenBank genomic sequence revealed two introns interrupting the H4 receptor sequence. The first intron was ~8 kb in length and located within the TM2-encoding region (interrupting the translated sequence "LNLAISDFFVG. . . VISIPLYIPH"). The second intron was downstream in the region encoding the second intracellular loop (interrupting the translated sequence "DRYLSVSNA. . . VSYRTQHTGV"). The length of this intron could not be determined, because the GenBank genomic sequence was incomplete.

A BLASTX search of the GenBank database with this novel DNA sequence revealed the greatest overall identity with the histamine receptor H3 (~40%) and, more distantly, with other amine-type receptors of the GPCR family (<30%). As also reported with the H3 receptor (Lovenberg et al., 1999), this novel receptor shared greater identity with other amine receptors including the serotonin, adrenergic, dopamine, and muscarinic receptors (~ 25%) than the histamine receptors H1 or H2 (~20%). An alignment with the known histamine receptors (Fig. 1) revealed a higher sequence similarity to the H3 receptor (58% in the TM regions) than to the H1 and H2 receptors (26 and 27% in the TM regions, respectively). It has been determined from their sequence similarities and phylogenetic analyses that the members of the histamine receptor family may have evolved from different ancestral genes and, through convergent evolution, acquired the residues to recognize and bind histamine (Leurs et al., 2000). This novel receptor, which we have named H4, seems to be a novel member of the histamine receptor subfamily, with closest relation to the H3 receptor. The amino acid sequence of the H4 receptor revealed many conserved residues and motifs found within the GPCR family (Fig. 2). Among these is an aspartic acid residue in TM 3 that is conserved in all cationic amine receptors and has been shown to be important in binding various amines to GPCRs (Savarese and Fraser, 1992). In addition, the human Met-TM2 probe was used to screen a rat genomic library, which retrieved a phage encoding the 5′ end of the novel gene from the start methionine to TM2. This fragment shared 71% identity (83% in the TM regions) to the human H4 sequence, revealing a rat H4 ortholog.

The human Met-TM2 and TM5-TM6 probes were used in Northern analyses of various human and rat tissues. In the rat, the TM5-TM6 probe revealed a single transcript of 3 kb

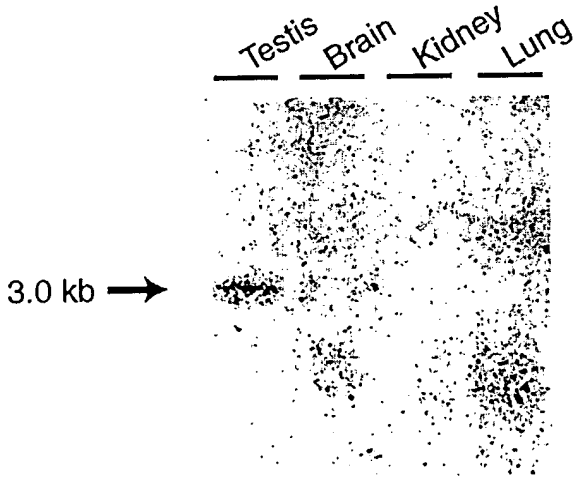


Fig. 3. Northern blot analysis of the distribution of H4 mRNA in various rat tissues. Each lane contained 10 μg of poly(A)$^+$ RNA.

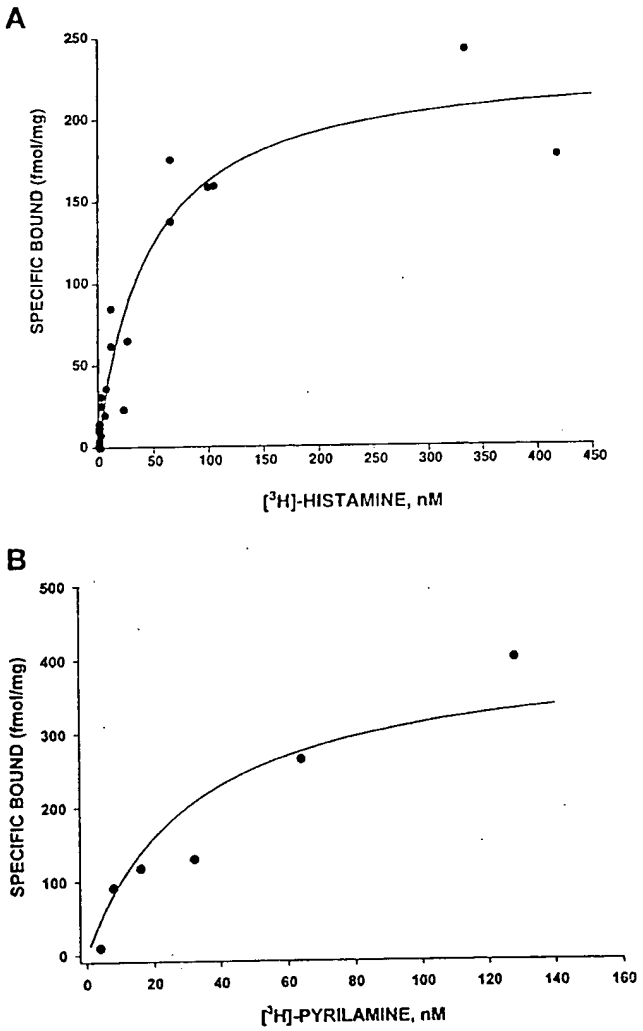


Fig. 4. Saturation binding studies of [^{3}H]histamine (A) and [^{3}H]pyrilamine (B) specific binding in membranes prepared from HEK-293 cells transiently transfected with H4 cDNA. Shown are typical results from experiment replicated three (A) and two times (B), with data representing mean of duplicate determinations. The curve represents the theoretical fit of the data A) (K_d = 44 nM; B_{max} = 235 fmol/mg) and B) (K_d = 32 nM; B_{max} = 437 fmol/mg). Nonspecific binding was determined with 100 μM mianserin.

in the testis (Fig. 3). The rat DNA fragment encoding from the start methionine to TM2 was used in Northern analyses of various rat tissues, revealing a 3-kb transcript in intestine (data not shown). The H3 and H4 receptors had significantly different mRNA expression distributions. H4 mRNA was detected in two peripheral tissues (with no detectable levels in brain or various peripheral tissues, including heart, stomach, small intestine, kidney, or liver). In contrast, H3 mRNA has been reported to be abundant in the brain (Lovenberg et al., 1999). Thus, H4 is not likely to be the H3-subtype characterized previously in brain tissue (West et al., 1990; Leurs et al., 1996), which suggests that yet another histamine receptor subtype remains unidentified.

To test the H4 receptor for pharmacological characterization, we constructed a full-length open reading frame by individually amplifying and joining the three H4-encoding fragments by PCR. The receptor was expressed in HEK-293 cells, and a variety of tritiated ligands including [^{3}H]histamine (histamine receptor nonselective), [^{3}H]pyrilamine (H1 receptor-selective), and [^{3}H]tiotidine (H2 receptor-selective) were tested against unlabeled ligands including mianserin, cyproheptadine, histamine, and clozapine. Specific binding obtained with [^{3}H]histamine (44 nM) or [^{3}H]pyrilamine (5–20 nM) and mianserin (100 μM) represented between 40 and 75% of total binding. Saturation binding studies performed with [^{3}H]histamine and [^{3}H]pyrilamine indicated respective K_d values of 44 and 32 nM and respective B_{max} values of 235 and 437 fmol/mg of membrane protein (Fig. 4). No specific [^{3}H]histamine or [^{3}H]pyrilamine binding was observed with untransfected HEK-293 cells.

Competition binding studies with various selective and nonselective histaminergic compounds were demonstrated. Typical data are shown in Fig. 5 and summarized in Table 1. The highest affinities were for amitriptyline and chlorpromazine, which are tricyclic compounds that have high affinity for the H1 histamine receptor (see on-line database at: *http://pdsp.cwru.edu/pdsp.asp*). Doxepin, cinnarizine, and promethazine (H1-selective antagonists) also displayed high affinity for the H4 receptor. Imetit (H3-selective agonist) and dimaprit (H2-selective agonist) had weak affinities for the H4 receptor, whereas mianserin (an H1 and H2 antagonist), cyproheptadine (a nonselective histamine/serotonin antagonist), and clozapine (an atypical antipsychotic drug with high affinity for a large number of receptors) had moderate affinities. The pharmacological profile of the H4 receptor is distinct from the histamine receptors (Table 1).

An HA epitope tag-encoding sequence was inserted after the start methionine for Western blot visualization. This plasmid was transiently transfected into COS-7 cells. Immunoblot analyses of membranes from these cells revealed high expression of the H4 receptor (Fig. 6A), with bands at 44 kDa, 85 kDa, and higher molecular mass species. The 44-kDa band matched the expected mass of the unglycosylated receptor and the 85-kDa band matched the glycosylated form. The bands >250 kDa represent oligomeric receptor species (Lee et al., 2000). The functional activity of the HA-tagged H4 receptor was examined by measuring histamine-induced internalization. As shown in Fig. 6B, exposure to 100 μM

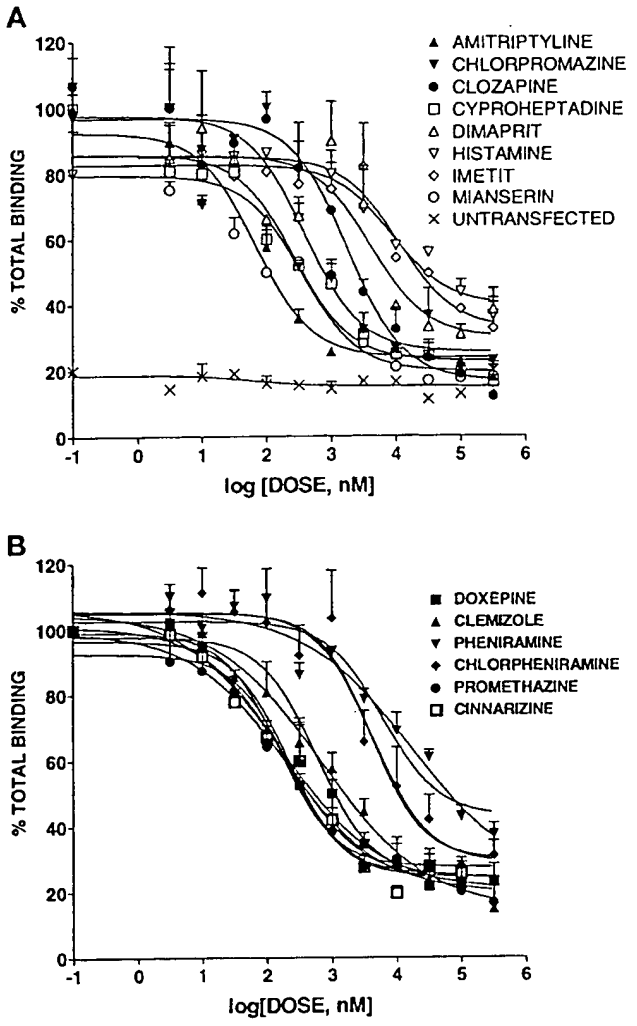


Fig. 5. Two sets (A and B) of competition binding of [^{3}H]pyrilamine with various ligands in membranes prepared from HEK-293 cells transiently transfected with H4 cDNA. Shown are typical results from competition binding experiments using 20 nM [^{3}H]pyrilamine and various concentrations of unlabeled ligands. Also shown are typical results obtained with untransfected HEK-293 cells studied in parallel (A).

TABLE 1

Ligand affinities for the H4 receptor

Data represent mean ± S.E. for two to four separate experiments. Twelve different concentrations spanning 6 log units of test drug were used to displace 15 nM [^{3}H]pyrilamine.

	$-\text{LogEC}_{50}$ (K_i)	Hill Coefficient
	nM	
Amitriptyline	7.31 ± 0.26 (33.6)	0.8
Chlorpromazine	7.13 ± 0.29 (50.2)	0.7
Doxepin	6.79 ± 0.13 (105.9)	0.6
Cinnarizine	6.73 ± 0.20 (141.6)	0.7
Promethazine	6.71 ± 0.19 (150.2)	0.6
Cyproheptadine	6.53 ± 0.09 (201.5)	0.9
Clemizole	6.16 ± 0.12 (402.2)	0.5
Mianserin	5.96 ± 0.40 (750.0)	0.9
Clozapine	5.90 ± 0.22 (849.6)	1.2
Chlorpheniramine	5.39 ± 0.16 (2910.0)	1.1
Histamine	5.30 ± 0.04 (3442.3)	0.5
Imetit	5.25 ± 0.44 (3795.6)	0.2
Pheniramine	5.21 ± 0.47 (4184.0)	0.2
Dimaprit	4.76 ± 0.19 (11812.7)	0.4
α-Methylhistamine	>10,000	Undetermined
Cimetidine	>10,000	Undetermined

histamine caused an internalization of H4 receptors from the plasma membrane to intracellular sites. Quantification revealed a rapid internalization of H4 receptors with significant internalization at 2 min after agonist exposure (Fig. 6C). These results indicated that the surface expression of the H4 receptor is functionally regulated by histamine exposure in a time-dependent fashion.

During the preparation of this article, other researchers (Oda et al., 2000) also reported the identification of a cDNA encoding a novel histamine receptor. This cDNA sequence varied from our sequence at three nucleotide positions, which translated into differences in three amino acids. Specifically, Ala138, His206, and Gln253 (as found in our sequence) were replaced by valine and two arginines, respectively. In the study by Oda et al., expression analyses revealed signals in peripheral blood leukocytes, small intestine, spleen and colon, and no expression in the brain. In addition, they reported histamine signaling through the novel receptor to be pertussis toxin-sensitive, suggesting a $G_{i/o}$ pathway of activation. For our receptor, we examined several second messenger-effector systems. We were not able to demonstrate H4 receptor-mediated inhibition of forskolin-stimulated adenylyl cyclase, alterations in phosphoinositide hydrolysis, or mitogen-activated protein kinase phosphorylation (extracellular signal-regulated kinase 1/2 phosphorylation) in HEK-293 cells.

In conclusion, we report the discovery of a novel histamine receptor, H4. Previously, it was observed that histamine receptors shared greater sequence similarities with other biogenic amine-binding GPCRs than with one another. H4 shared highest sequence similarity with the previously reported histamine H3 receptor. In combination with the H1, H2, and H3 receptors, this receptor, with its unique distribution and pharmacology, will undoubtedly provide further insight into the physiological functions and therapeutic applications of this receptor family.

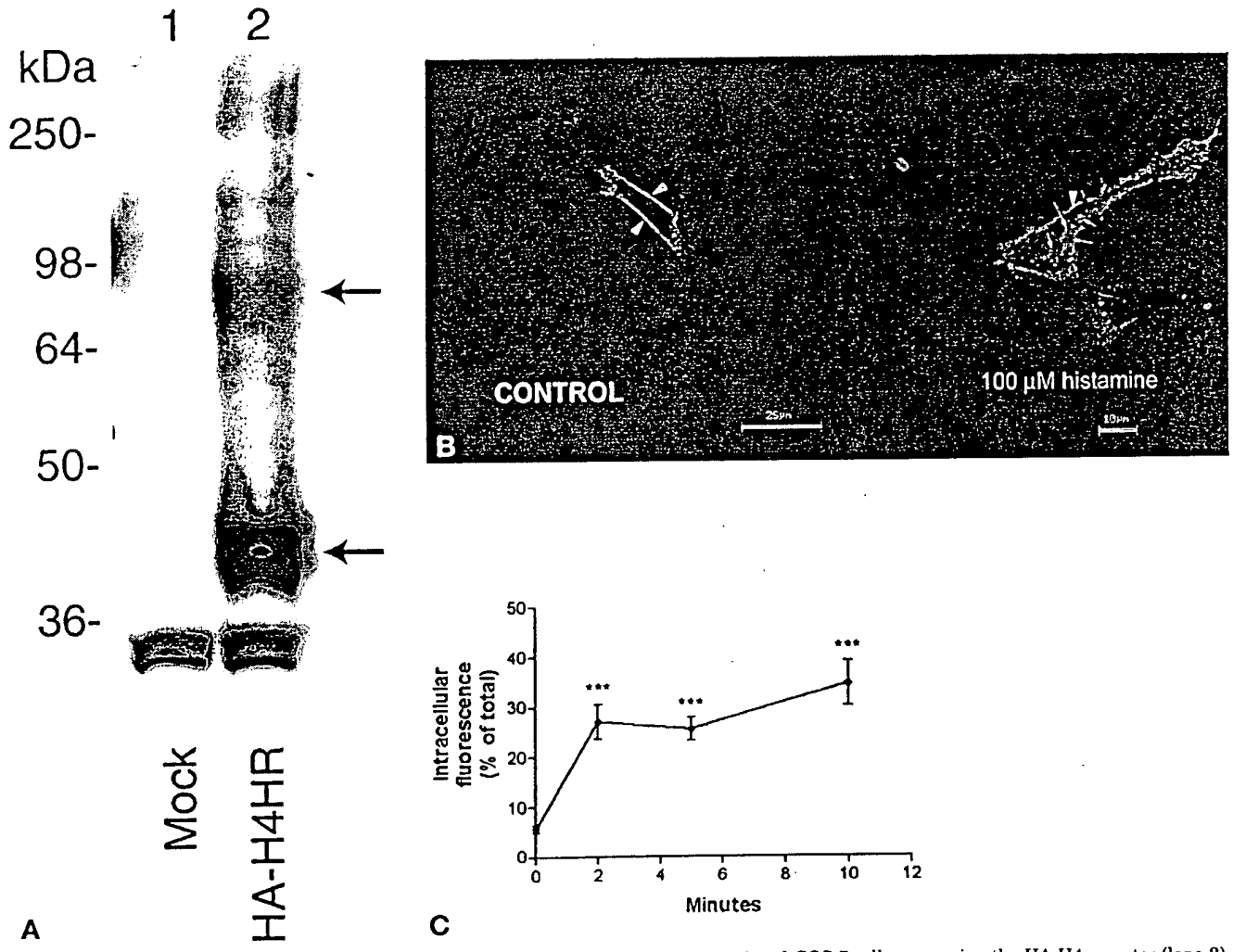


Fig. 6. A, immunoblot analysis of membranes from mock-transfected COS-7 cells (lane 1) and COS-7 cells expressing the HA-H4 receptor (lane 2). Membrane protein (25 μg) was used in each lane. Arrows indicate the unglycosylated (bottom) and glycosylated (top) receptor. B) histamine-induced internalization of the H4 receptor in HEK-293 cells. Shown are representative confocal micrographs in which HA-tagged H4 receptors were examined in cells exposed to vehicle (PBS) or 100 μM histamine for 5 min and then prepared for microscopy as described previously (Berry et al., 1996; Willins et al., 1999; Kristiansen et al., 2000). Arrowheads and arrows indicate cell-surface and internalized receptors, respectively. C, histamine-induced time-dependent internalization of the H4 receptor in HEK-293 cells. Shown are the mean ± S.E.M. of the percentage internalization of HA-tagged H4 receptors in HEK-293 cells (n = 20–30 cells/time point) after exposure to histamine (100 μM) for various time periods. ***significantly different compared with 0-min value ($P < 0.0001$).

References

Berry SA, Shah MC, Khan N and Roth BL (1996) Rapid agonist-induced internalization of the 5-hydroxytryptamine2A receptor occurs via the endosome pathway in vitro. *Mol Pharmacol* 50:306–313.

Gantz I, Schaffer M, DelValle J, Logsdon C, Campbell V, Uhler M and Yamada T (1991) Molecular cloning of a gene encoding the histamine H2 receptor. *Proc Natl Acad Sci USA* 88:429–433.

Heise CE, O'Dowd BF, Figueroa DJ, Sawyer N, Nguyen T, Im DS, Stocco R, Bellefeuille JN, Abramovitz M, Cheng R, et al. (2000) Characterization of the human cysteinyl leukotriene 2 receptor. *J Biol Chem* 275:30531–30536.

Hill SJ, Ganellin CR, Timmerman H, Schwartz JC, Shankley NP, Young JM, Schunack W, Levi R and Haas HL (1997) International Union of Pharmacology. XIII. Classification of histamine receptors. *Pharmacol Rev* 49:253–278.

Kolakowski LF Jr, O'Neill GP, Howard AD, Broussard SR, Sullivan KA, Feighner SD, Sawzdargo M, Nguyen T, Kargman S, Shiao LL, et al. (1998) Molecular characterization and expression of cloned human galanin receptors GALR2 and GALR3. *J Neurochem* 71:2239–2251.

Kristiansen K, Kroeze WK, Willins DL, Gelber EI, Savage JE, Glennon RA and Roth BL (2000) A highly conserved aspartic acid (Asp-155) anchors the terminal amine moiety of tryptamines and is involved in membrane targeting of the 5-HT(2A) serotonin receptor but does not participate in activation via a "salt-bridge disruption mechanism. *J Pharmacol Exp Ther* 293:735–746.

Lee DK, George SR, Evans JF, Lynch KR and O'Dowd BF (2001) Orphan G protein-coupled receptors in the CNS. *Curr Opin Pharmacol*, in press.

Lee SP, O'Dowd BF, Ng GY, Varghese G, Akil H, Mansour A, Nguyen T and George SR (2000) Inhibition of cell surface expression by mutant receptors demonstrates that D2 dopamine receptors exist as oligomers in the cell. *Mol Pharmacol* 58:120–128.

Leurs R, Hoffmann M, Wieland K and Timmerman H (2000) H3 receptor gene is cloned at last. *Trends Pharmacol Sci* 21:11–12.

Leurs R, Kathmann M, Vollinga RC, Menge WM, Schlicker E and Timmerman H (1996) Histamine homologues discriminating between two functional H3 receptor assays. Evidence for H3 receptor heterogeneity? *J Pharmacol Exp Ther* 276:1009–1015.

Lovenberg TW, Roland BL, Wilson SJ, Jiang X, Pyati J, Huvar A, Jackson MR and Erlander MG (1999) Cloning and functional expression of the human histamine H3 receptor. *Mol Pharmacol* 55:1101–1107.

Marchese A, Docherty JM, Nguyen T, Heiber M, Cheng R, Heng HH, Tsui LC, Shi X, George SR and O'Dowd BF (1994) Cloning of human genes encoding novel G protein-coupled receptors. *Genomics* 23:609–618.

Marchese A, George SR, Kolakowski LF Jr, Lynch KR and O'Dowd BF (1999) Novel GPCRs and their endogenous ligands: Expanding the boundaries of physiology and pharmacology. *Trends Pharmacol Sci* 20:370–375.

Ng GY, O'Dowd BF, Lee SP, Chung HT, Brann MR, Seeman P and George SR (1996) Dopamine D2 receptor dimers and receptor-blocking peptides. *Biochem Biophys Res Commun* 227:200–204.

Nguyen T, Erb L, Weisman GA, Marchese A, Heng HH, Garrad RC, George SR, Turner JT and O'Dowd BF (1995) Cloning, expression, and chromosomal localization of the human uridine nucleotide receptor gene. *J Biol Chem* 270:30845–30848.

Oda T, Morikawa N, Saito Y, Masuho Y and Matsumoto SI (2000) Molecular cloning and characterization of novel type of histamine receptor preferentially expressed in leukocytes. *J Biol Chem* 275:36781–36786.

O'Dowd BF, Lee DK, Huang W, Nguyen T, Cheng R, Liu Y, Wang B, Gershengorn MC and George SR (2000) TRH-R2 exhibits similar binding and acute signaling but distinct regulation and anatomic distribution compared with TRH-R1. *Mol Endocrinol* 14:183–193.

Retief JD, Lynch KR and Pearson WR (1999) Panning for genes—A visual strategy for identifying novel gene orthologs and paralogs. *Genome Res* 9:373–382.

Savarese TM and Fraser CM (1992) In vitro mutagenesis and the search for structure-function relationships among G protein-coupled receptors. *Biochem J* 283:1–19.

Shapiro DA, Kristiansen K, Kroeze WK and Roth BL (2000) Differential modes of agonist binding to 5-hydroxytryptamine$_{2A}$ serotonin receptors revealed by mutation and molecular modeling of conserved residues in transmembrane region 5. *Mol Pharmacol* 58:877–886.

West RE Jr, Zweig A, Shih NY, Siegel MI, Egan RW and Clark MA (1990) Identification of two H3-histamine receptor subtypes. *Mol Pharmacol* 38:610–613.

Willins DL, Berry SA, Alsayegh L, Backstrom JR, Sanders-Bush E, Friedman L and Roth BL (1999) Clozapine and other 5-hydroxytryptamine-2A receptor antagonists alter the subcellular distribution of 5-hydroxytryptamine-2A receptors in vitro and in vivo. *Neuroscience* 91:599–606.

Yamashita M, Fukui H, Sugama K, Horio Y, Ito S, Mizuguchi H and Wada H (1991) Expression cloning of a cDNA encoding the bovine histamine H1 receptor. *Proc Natl Acad Sci USA* 88:11515–11519.

Send reprint requests to: Dr. Brian F. O'Dowd, Department of Pharmacology, University of Toronto, Medical Science Building, 8 Taddle Creek Rd. Rm 4353, Toronto, Ontario, Canada M5S 1A8. E-mail: brian.odowd@utoronto.ca

In re: Meyers
Appl. No. 09/464,039
Filed December 15, 1999

APPENDIX G

CANCER GENETICS

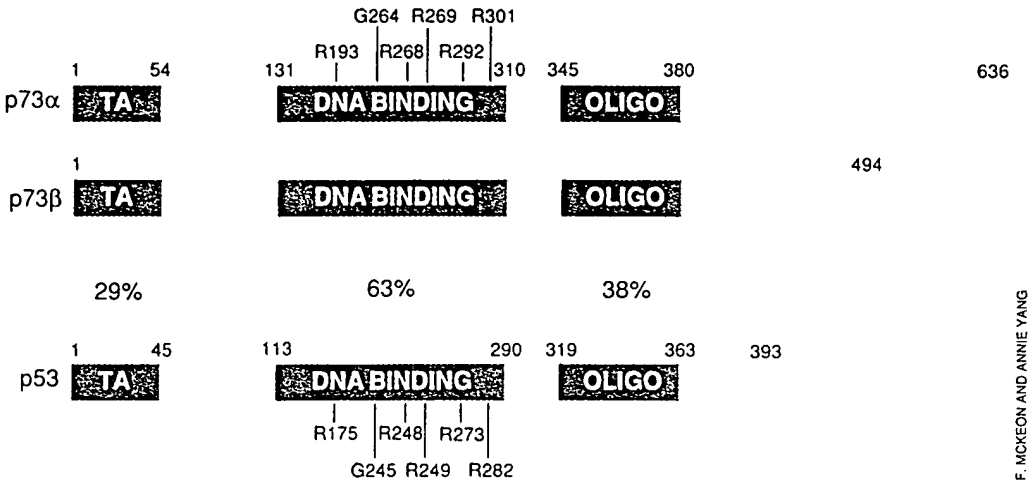
First *p53* Relative May Be a New Tumor Suppressor

In the 20 years since its discovery, the *p53* gene has become one of the most heavily scrutinized genes in history. Indeed, it's referenced in over 8000 papers in Medline, the online biomedical abstract service. The fascination is easy to understand: Loss or inactivation of *p53*, which is a so-called tumor-suppressor gene, is thought to contribute to the development of 50% of all human cancers. All that time, *p53* was thought to be an only child, with no close relatives. Now, researchers have discovered a new gene, a long-lost cousin called *p73*, that bears a strong resemblance to *p53*.

It is being greeted with the same surprise as any newfound relative. "Given the intense interest in this area, the fact that [*p73*] slipped through the cracks is surprising," says cancer biologist Tyler Jacks of the Massachusetts Institute of Technology (MIT). But the new gene should generate some intense interest of its own, because its protein not only resembles the p53 protein, but also seems to have similar activities. The p53 protein acts as a "security guard," deployed when a cell's DNA is damaged to prevent the cell from becoming cancerous. It does this by either inhibiting cell growth until the damage is repaired or causing the cell to commit suicide through a process called programmed cell death or apoptosis. The p73 protein appears to share these growth-inhibiting and apoptosis-promoting effects, although what triggers them and exactly what its cellular role is are both unknown.

Those findings, together with *p73*'s location in a region of chromosome 1 that is often deleted in cancers including neuroblastoma, a malignant tumor of nervous tissue, suggest that it, too, may be a tumor suppressor. "This [discovery] will titillate a whole lot of people," predicts cancer geneticist Bert Vogelstein of Johns Hopkins University School of Medicine, a pioneer of *p53* research. Indeed, if *p73* can stand in for *p53* when that gene is lost, it might be possible to design new cancer drugs that work by turning on *p73* in tumors lacking *p53*.

Molecular biologist Daniel Caput and his colleagues at the pharmaceutical company Sanofi Recherche in Labege, France, identified the *p73* gene while looking for something completely different, namely genes that respond to certain immune system regulators. When the French team sequenced the many potential targets their screen had turned up, they were shocked to find that one false positive had remarkable similarities to *p53*.



F. MCKEON AND ANNIE YANG

No longer alone. The p73 proteins, although longer, resemble p53 in three regions: the transcription activation (TA, 29% identical) and DNA binding domains (63% identical) and also the domain where p53 binds itself (OLIGO, 38% identical). The labeled amino acids indicate residues that are frequently mutated in *p53* and are conserved in *p73*.

As the researchers report in the 22 August issue of *Cell*, the proteins made by *p73* are somewhat larger than p53. But they found that one section of p73 closely resembles the so-called "core binding region" of p53. Many of p53's activities depend on its ability to regulate other genes, and the core binding region is where the protein attaches itself to the DNA when exerting its effects. Of 177 amino acids in that region, 112 are identical. Additional similarities turned up in two other sections thought to be involved in p53 activity—one needed for its gene regulatory effects and another where it apparently binds to itself. These resemblances are enough to suggest that the two genes are the progeny of a gene that was duplicated in some ancient cellular event. Indeed, *p73* may be the ancestral gene, because a gene found in squid that was supposed to be that species's version of *p53* is actually more similar to *p73*.

"This [discovery] will titillate a whole lot of people."

—Bert Vogelstein

The structural similarities between p53 and p73 also suggested that the proteins might have similar roles in the cell. So Caput and his colleagues joined forces with their longtime collaborator, cell biologist Frank McKeon, who studies gene expression and cell division at Harvard Medical School in Boston, to look for parallels. One way p53 restrains cells that have damaged DNA is by triggering the production of a protein called p21, which blocks cell division. The Caput-McKeon team found that adding *p73* to a line of neuroblastoma that lacks the gene also triggers p21 production, an indication that p73 inhibits cell growth through the same pathway used by p53.

In a paper that appears in this week's issue of *Nature*, molecular biologist William Kaelin at the Dana-Farber Cancer Institute in Boston and his colleagues report similar findings with another tumor cell line. Kaelin's team also found evidence that p73 can mimic p53's ability to cause cell suicide. When overexpressed in these cells, p73 latched onto stretches of DNA to which p53 normally attaches itself when instructing a cell to self-destruct.

Together, the findings suggest that *p73*, too, may be a tumor suppressor, an idea that is buttressed by its provocative chromosomal location. The Caput-McKeon team found *p73* in a region near the tip of chromosome 1 that was already suspected of harboring one or more tumor suppressor genes, because the region is often missing in tumor cells.

The teams did find one major point of difference between the two genes, however. Unlike p53 protein, p73 does not seem to be produced in response to DNA damage. That implies that the protein is not a cell "security guard" the way p53 is. Early results of experi-

ments in which McKeon and Caput deleted the *p73* gene in mice suggest another possibility: It may be "developmentally important," he says, especially in the brain and immune system, although how remains to be clarified.

If *p73* is a tumor suppressor, it may behave somewhat differently than *p53* and other previously discovered tumor suppressors. Classic tumor-suppressor genes require two "hits" to be inactivated—a partial or complete deletion of one of the two gene copies, for example, and another, lesser change that cripples the second copy. But Caput, McKeon, and their colleagues have evidence that one *p73* copy is already inactive in normal cells—the apparent result of a mysterious process called imprinting. Its precise function isn't known, but during embryonic development, imprinting alters certain genes so that the copy inherited either from the mother or the father is specifically shut down.

If one *p73* copy has been silenced by imprinting, then only one hit—loss of the active copy—might be all that it takes to tip a cell into the uncontrolled growth of cancer. Says Kaelin, "*p73* may be the first example of a new paradigm for how tumor-suppressor genes are involved in cancer."

Indeed, molecular biologist Rogier Versteeg of the Academic Medical Center in Amsterdam, the Netherlands, has evidence that an imprinted gene may be involved in neuroblastoma development. He has identified two sites of chromosome damage that contribute to neuroblastoma by knocking out as-yet-undiscovered tumor-suppressor genes. Both lie in the same region of chromosome 1 where *p73* is located, and one illustrates "a strong bias" toward loss from the maternal copy of the chromosome in the cancer cells. This bias implies that this specific copy is the active one and must be lost to cause the cancer.

Other work from the Caput-McKeon team suggests that this mystery tumor-suppressor gene may be *p73*. When they looked for the gene in neuroblastoma cell lines, they found that one *p73* copy had been lost. And while they couldn't uncover any mutations in the remaining copy, most of the cell lines made no detectable p73 protein, implying that the second copy had been silenced by imprinting.

In spite of the differences in the roles of *p53* and its new cousin, both in normal cells and in cancer, the family resemblances may be strong enough for them to substitute for each other. If so, says MIT's Jacks, cancer might be treated by finding a way to switch on *p73* in tumor cells that have lost *p53*. "Even if *p73* is not normally involved in tumor suppression, maybe it could be recruited," says Jacks. Now McKeon and Caput are searching for further family members. But the discovery of *p73* is already certain to captivate their peers.

–Steven Dickman

Steven Dickman is a free-lance writer in Cambridge, Massachusetts.

AIDS RESEARCH

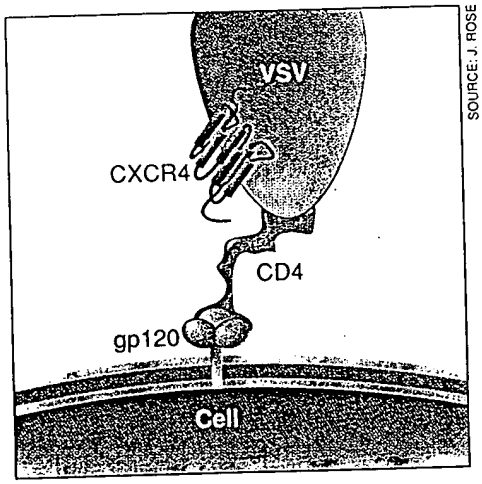
HIV Gets a Taste of Its Own Medicine

In an attempt to fight fire with fire, researchers have engineered a virus that usually infects cattle to attack the AIDS virus in humans. The innovative approach has so far shown promise only in test-tube experiments, but it is attracting widespread attention among AIDS researchers. "It's really on the verge of a breakthrough," says Nava Sarver, who oversees development of novel AIDS treatments at the National Institute of Allergy and Infectious Diseases (NIAID).

Yale University virologist John Rose and co-workers describe in the 5 September issue of *Cell* how they have constructed a potential HIV treatment by modifying vesicular stomatitis virus (VSV), which farmers detest because it causes a mouth infection in cattle that prevents them from eating. As the Yale researchers' experiments show, their newfangled VSV selectively targets and destroys HIV-infected human cells. "It's a pretty interesting way of harnessing a virus for peaceful purposes," says the University of Pennsylvania's Robert Doms. "It's a very clever approach."

The work builds on recent discoveries made by Doms and others about a two-part handshake between HIV and the cells it infects. After HIV binds to the CD4 receptor on a white blood cell, it also must link to another molecule found on the cell's surface, known as a chemokine receptor. Once these handshakes are complete, HIV gains entry, and shortly thereafter, new virus proteins make their way to the cell's outer coating, where they stick out like a flag of victory.

Rose and colleagues reasoned that if VSV could be induced to express these receptors on its surface, they would bind to the HIV proteins displayed on infected cells, turning VSV into a kind of guided missile. To test this idea, the researchers stitched into VSV the genes that code for CD4 and one of HIV's favored chemokine receptors, CXCR4, and added their engineered VSV to a culture containing HIV-infected cells. The virus did, indeed, target just the infected cells, killing them rapidly. "VSV is so fast," says Rose—much faster than HIV, he notes.



SOURCE: J. ROSE

Trojan horse. CD4 and CXCR4 receptors expressed by genetically engineered VSV bind to HIV's gp120 protein on surface of infected cell.

A potential downside to this approach is that the modified VSV might kill cells that aren't infected by HIV. Rose believes that won't happen because he has stripped VSV of its own surface protein, which is what allows it to infect a broad range of cells. "Without its normal coat, it can't infect anything," says Rose. But only animal tests will provide evidence of that, cautions NIAID director Anthony Fauci.

Although Fauci has high praise for the concept's ingenuity, he is concerned that it might take an impracticably high dose of the modified VSV to make a real dent in a person's HIV levels. Another worry, says monkey researcher Ronald Desrosiers of the New England Regional Primate Research Center in Southborough, Massachusetts, is that the body will quickly develop an immune response against VSV, limiting its ability to attack HIV.

Still, Sarver, Fauci, and others are anxious for Rose and colleagues to put their viral guided missile to more stringent test-tube and animal tests. Desrosiers already has begun working with Rose to test the concept in monkeys that have been infected with SIV, HIV's simian cousin. Desrosiers expects to have results in the next few months. Even if they are positive, however, human trials will require the approval of the Food and Drug Administration, which has shown great caution in the past about putting potentially therapeutic viruses into people.

Rose's strategy is not limited to attacking HIV. NIAID's Sarver suggests that if researchers can swap different receptors into this "gutted" VSV, the precisely targeted viruses could be used in everything from vaccines to gene therapies to cancer treatments. "We're not there yet," says Sarver, "but the potential applications are enormous."

–Jon Cohen

In re: Meyers
Appl. No. 09/464,039
Filed December 15, 1999

APPENDIX H

Cell, Vol. 92, 73–82, January 9, 1998

An Orphan Nuclear Receptor Activated by Pregnanes Defines a Novel Steroid Signaling Pathway

Steven A. Kliewer,*# John T. Moore,†
Laura Wade,* Jeff L. Staudinger,*
Michael A. Watson,* Stacey A. Jones,*
David D. McKee,† Beverly B. Oliver,*
Timothy M. Willson,‡ Rolf H. Zetterström,§
Thomas Perlmann,‖ and Jürgen M. Lehmann*
*Department of Molecular Endocrinology
†Department of Molecular Sciences
‡Department of Medicinal Chemistry
Glaxo Wellcome Research and Development
Research Triangle Park, North Carolina 27709
§Department of Neuroscience
Karolinska Institute
S-171 77 Stockholm
Sweden
‖The Ludwig Institute for Cancer Research
Stockholm Branch
S-171 77 Stockholm
Sweden

Summary

Steroid hormones exert profound effects on differentiation, development, and homeostasis in higher eukaryotes through interactions with nuclear receptors. We describe a novel orphan nuclear receptor, termed the pregnane X receptor (PXR), that is activated by naturally occurring steroids such as pregnenolone and progesterone, and synthetic glucocorticoids and antiglucocorticoids. PXR exists as two isoforms, PXR.1 and PXR.2, that are differentially activated by steroids. Notably, PXR.1 is efficaciously activated by pregnenolone 16α-carbonitrile, a glucocorticoid receptor antagonist that induces the expression of the CYP3A family of steroid hydroxylases and modulates sterol and bile acid biosynthesis in vivo. Our results provide evidence for the existence of a novel steroid hormone signaling pathway with potential implications in the regulation of steroid hormone and sterol homeostasis.

Introduction

Steroid hormones are cholesterol derivatives that serve as signaling molecules to coordinate the expression of complex gene programs in higher eukaryotes. These molecules exert their effects by diffusing into cells and interacting with specific intracellular receptors. Receptors for each of the major classes of sex and adrenal steroids have been characterized (Evans, 1988; Beato et al., 1995; Mangelsdorf et al., 1995). In the absence of their cognate ligands, the steroid hormone receptors remain sequestered in the cytoplasm through interactions with large multiprotein complexes containing heat shock proteins. However, the binding of ligand causes the steroid hormone receptors to be released from these complexes and translocated into the nucleus (Pratt, 1993). Once inside the nucleus, the activated receptors regulate the expression of target genes by binding as homodimers to short DNA sequence motifs, termed hormone response elements (HREs) (Glass, 1994). In this manner, the steroid hormone receptors function as ligand-activated transcription factors.

The molecular cloning of steroid hormone receptors revealed that they comprise a subfamily within a larger superfamily of structurally related proteins (Evans, 1988; Mangelsdorf and Evans, 1995; Mangelsdorf et al., 1995). This superfamily also includes receptors for nonsteroidal, lipophilic molecules such as thyroid hormone, retinoids, fatty acids, and eicosanoids. The nonsteroid receptors differ from their steroid hormone receptor counterparts in several respects (Mangelsdorf and Evans, 1995; Mangelsdorf et al., 1995). First, the nonsteroid hormone receptors are not sequestered in the cytoplasm in the absence of their cognate ligands but instead reside within the nucleus. Second, whereas steroid hormone receptors generally bind to their HREs as homodimers, most of the nonsteroid hormone receptors identified to date bind to DNA as heterodimers with the 9-*cis* retinoic acid receptors (RXRs) (Glass, 1994; Mangelsdorf and Evans, 1995). Finally, the two classes of receptors recognize different types of HREs: steroid hormone receptors generally bind to HREs composed of two half-sites organized as a palindrome with a three-nucleotide spacer, while nonsteroid hormone receptors preferentially bind to HREs composed of two half-sites organized as a direct repeat (DR), with the number and composition of the nucleotides separating the half-sites serving as important determinants of receptor selectivity (Umesono et al., 1991; Glass, 1994; Mangelsdorf and Evans, 1995).

In addition to receptors with established ligands, approximately 30 other members of the nuclear hormone receptor family have been isolated from vertebrates. These related proteins lack identified ligands and, as a consequence, have been termed orphan nuclear receptors (Evans, 1988; Mangelsdorf and Evans, 1995; Enmark and Gustafsson, 1996). The search for hormonal activators of the orphan receptors has created exciting opportunities to discover novel endocrine signaling pathways with implications in both normal physiology and disease. Work with orphan nuclear receptors has led to the identification of fatty acids and eicosanoids as ligands for the peroxisome proliferator–activated receptors (Forman et al., 1997; Kliewer et al., 1997; Krey et al., 1997), retinoids and farnesoids as activators of RIP14/FXR (Forman et al., 1995; Zavacki et al., 1997), and various oxysterols as activators of the LXR and SF-1 orphan nuclear receptors (Janowski et al., 1996; Lala et al., 1997; Lehmann et al., 1997).

In this report, we describe a novel orphan nuclear receptor, which we have designated PXR for pregnane X receptor. Like nonsteroid hormone receptors, PXR binds as a heterodimer with RXR to an HRE composed of two half-sites organized as a DR. Surprisingly, however,

To whom correspondence should be addressed.

PXR is efficaciously activated by several steroids, including naturally occurring pregnanes and synthetic glucocorticoids and antiglucocorticoids. Thus, PXR combines features of both the steroid and nonsteroid nuclear receptors. We suggest that PXR defines a novel steroid hormone signaling pathway that may account for at least some of the effects of synthetic glucocorticoids and antiglucocorticoids that do not appear to be mediated through the classical glucocorticoid receptor (GR) signaling pathway.

Results

Cloning of PXR.1 and PXR.2

In an effort to identify new members of the nuclear receptor family, we performed a series of motif searches of public EST databases. These searches revealed a clone from a mouse liver library in the Washington University/HHMI EST database that had homology to the ligand-binding domains (LBDs) of a number of nuclear receptors. We used this partial sequence information to isolate larger clones from a mouse liver cDNA library. The nucleotide sequence of the longest cDNA clone encodes a novel orphan nuclear receptor of 431 amino acids that we have designated PXR.1 (Figure 1A). We also isolated a second cDNA clone, termed *PXR.2*, that was identical to *PXR.1* except for the deletion of a stretch of 123 nucleotides extending from base pairs 661 to 783 (Figure 1A). Examination of the *PXR* genomic structure revealed that *PXR.2* represents a splice variant of *PXR.1* lacking a single exon (data not shown). The *PXR.2* cDNA encodes a 390 amino acid protein that lacks a 41 amino acid region in the putative LBD of PXR.1. Sequence alignment with nuclear receptors for which the crystal structures have been solved indicates that the 41 amino acids that distinguish the PXR isoforms lie between helices two and three of the canonical LBD structure (Wurtz et al., 1996).

Sequence comparison with other members of the nuclear receptor family showed that the PXR isoforms are most closely related to the *Xenopus laevis* orphan nuclear receptor 1 (ONR1) (Smith et al., 1994), with PXR.2 and ONR1 sharing 70% and 46% amino acid identity in their DNA-binding domains (DBDs) and LBDs, respectively (Figure 1B). Based upon this degree of homology, it is unclear at present whether PXR represents the mammalian homolog of ONR1. With regard to mammalian receptors, PXR is most closely related to the vitamin D receptor (VDR) (Baker et al., 1988), sharing 64% and 39% identity in the DBD and LBD, respectively.

PXR Expression Pattern in the Embryo and the Adult

The expression pattern of PXR was examined in both adult and embryonic tissues. Northern blot analysis was performed under high-stringency conditions with blots that included poly(A)$^+$ RNA prepared from multiple adult mouse tissues and a probe that recognized both PXR isoforms. Abundant expression of *PXR* mRNA was observed only in the liver and intestine, where three distinct messages were detected including a highly expressed mRNA species of 2.6 kb and two transcripts of 1.9 kb and 4.4 kb that were expressed at lower levels (Figure 2A). Weaker expression of the *PXR* mRNA was also detected in kidney and stomach (Figure 2A). No *PXR* mRNA was detected in the other tissues examined.

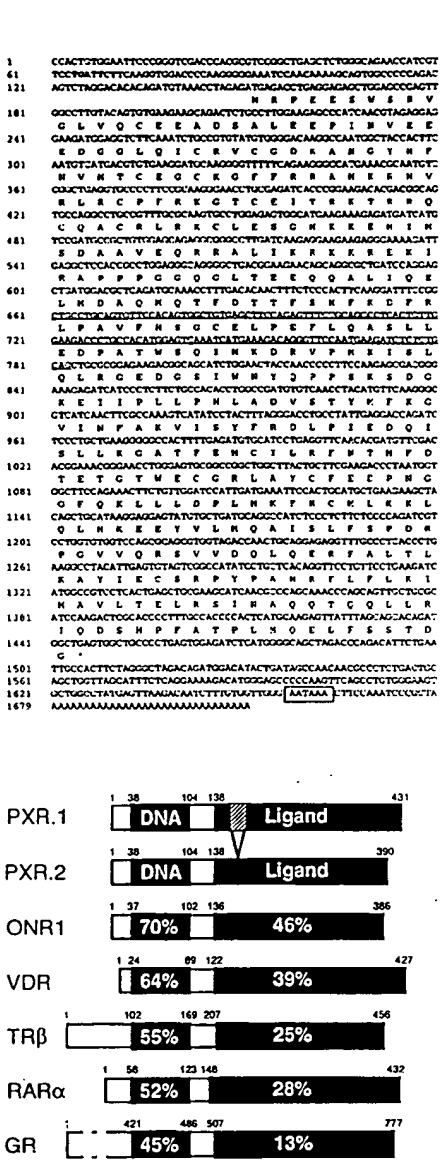


Figure 1. PXR is a Member of the Nuclear Receptor Superfamily

(A) Nucleotide and predicted amino acid sequence of mouse PXR. The upstream in-frame stop codon is in bold. The predicted start initiation codon is at nucleotide 151. The 123 nucleotide region extending from base pairs 661 to 783 that is present in PXR.1 and absent in PXR.2 is underlined. A putative polyadenylation signal in the 3′ untranslated region is boxed.

(B) Amino acid sequence comparison between murine PXR and other members of the nuclear hormone receptor family. Similarity between PXR and other nuclear hormone receptor family members in the DNA- and ligand-binding domains are indicated as percent amino acid identity. The 41 amino acid region that distinguishes PXR.1 from PXR.2 is indicated by the cross-hatched area. All comparisons were made with the PXR.2 isoform. ONR1, *Xenopus* orphan nuclear receptor 1; VDR, human vitamin D receptor; TRβ, human thyroid hormone receptor β; RARα, human retinoic acid receptor α; GR, human glucocorticoid receptor.

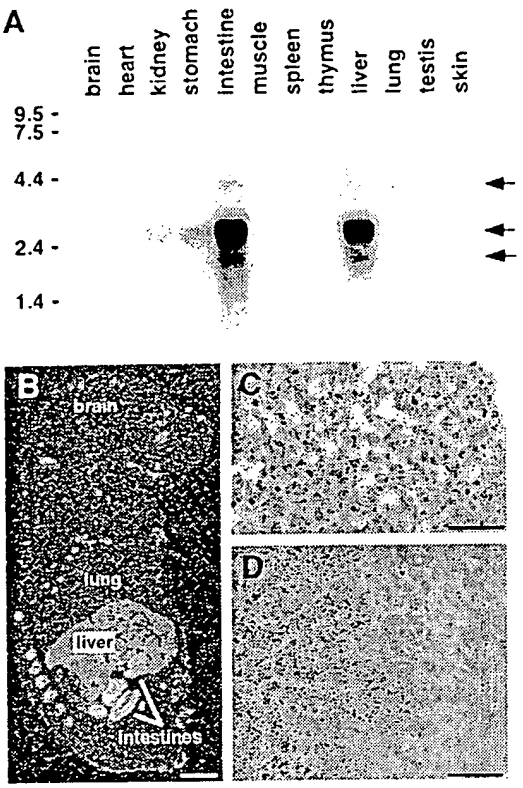


Figure 2. Expression Pattern of PXR in Adult and Embryonic Tissues

(A) Northern blot analysis. RNA size markers (in kb) are indicated at left. Arrows at right indicate the three species of *PXR* transcripts that were detected. (B-D) In situ hybridization analysis of *PXR* mRNA expression in E18 mouse embryo sections with a probe that recognized both PXR isoforms (see probe 1 in Experimental Procedures). (B) Phosphoimage of an entire sagittal section showing specific labeling in intestine and liver. Specificity was ascertained by competition with a 100-fold excess of unlabeled, specific oligonucleotide. This resulted in complete inhibition of labeling in liver and intestine while nonspecific signals remained in developing bone. (C and D) High power bright-field microscopy showing the abundance of silver grains in the liver (C) and in the intestine (D) where mRNA labeling was confined to the epithelium. The bar in (B) corresponds to 2 mM, and in (C) and (D) to 50 μM.

The embryonic expression pattern of *PXR* was examined via in situ hybridization analysis using sections prepared from day 18 (E18) mouse embryos. Three oligonucleotide probes were designed that recognized *PXR*, including two that interacted with both *PXR* isoform mRNAs and a third that hybridized only to *PXR.1* mRNA. *PXR* mRNA was detected in the liver and intestine (Figures 2B–2D). Staining in the intestine was confined to the epithelium (Figure 2D). No *PXR* expression was detected in embryonic kidney, adrenal, lung, thymus, heart, skeletal muscle, brain, or spinal cord (data not shown). While similar results were obtained with all three *PXR* probes, the signal intensity in liver and intestine was consistently higher in experiments performed with the probes that recognize both *PXR* isoforms (data not shown), suggesting that *PXR.2* is expressed in tissues that also express *PXR.1*. We conclude from the Northern blot and in situ hybridization analyses that *PXR* is abundantly expressed in only a few tissues, including the liver and intestine, in both the mouse embryo and adult.

PXR Is Activated by Synthetic Pregnanes and Glucocorticoids

We next sought to determine whether PXR, like other members of the nuclear receptor family, possesses transcriptional activity that can be regulated in a hormone-dependent manner. As we lacked knowledge of a cognate HRE for PXR, we initially performed searches for activators using an established chimera system in which the LBDs of the two PXR isoforms were fused to the DBD of the yeast transcription factor GAL4 (Lehmann et al., 1995). Expression vectors for the GAL4-PXR chimeras were transiently transfected into CV-1 cells together with a reporter plasmid containing five copies of a GAL4 DNA binding site upstream of the chloramphenicol acetyltransferase (CAT) reporter. Transfected CV-1 cells were systematically treated with a series of natural and synthetic compounds that included steroids, vitamin D analogs, thyroid hormone analogs, retinoids, fatty acids, and other small, lipophilic molecules, and reporter levels measured.

Interestingly, we found that the activity of the GAL4-PXR.1 chimera was markedly induced by 10 μM concentrations of a variety of synthetic steroids including the glucocorticoids dexamethasone, dexamethasone-t-butyl-acetate, and dexamethasone-21-acetate, and the pregnenolone derivative 6,16α-dimethyl pregnenolone (Figure 3B). Dexamethasone-t-butylacetate and 6,16α-dimethyl pregnenolone were the most efficacious of these compounds (Figure 3B). Remarkably, we found that PXR.1 was not only activated by GR agonists, but also by the GR antagonists RU486 and pregnenolone 16α-carbonitrile (PCN) (Figure 3B). RU486 is a potent antiprogestin that binds to the GR at nanomolar concentrations (Cadepond et al., 1997). PCN is a weaker antagonist that has been shown to interfere with GR-mediated activation at micromolar concentrations (Schuetz and Guzelian, 1984; Schuetz et al., 1984). Thus, the LBD of PXR.1 was efficaciously activated by both agonists and antagonists of the GR.

Notably, the GAL4-PXR.2 chimera displayed a much more restricted activation profile. Of the steroids that activated GAL4-PXR.1, only dexamethasone-t-butyl-acetate activated GAL4-PXR.2 efficiently (Figure 3B). We conclude that the 41 amino acid deletion that distinguishes PXR.2 from PXR.1 has a marked effect on the responsiveness of the orphan receptor to synthetic steroids.

PXR Functions through a Response Element Conserved in the *CYP3A* Gene Promoters

PCN and dexamethasone treatment have previously been shown to induce the expression of the *CYP3A* family of genes in rodent liver, intestine, and kidney as well as in primary cultures of rodent hepatocytes (Elshourbagy and Guzelian, 1980; Heuman et al., 1982; Hardwick et al., 1983; Schuetz and Guzelian, 1984; Schuetz et al., 1984; Gonzalez et al., 1985; Debri et al., 1995). The *CYP3A* genes encode cytochrome P450 hemoproteins involved in the hydroxylation of steroid hormones, including corticosteroids, progestins, androgens, and DHEA-sulfate, as well as a variety of xenobiotics (Nebert and Gonzalez, 1987; Juchau, 1990). The response to PCN and dexamethasone occurs at the level

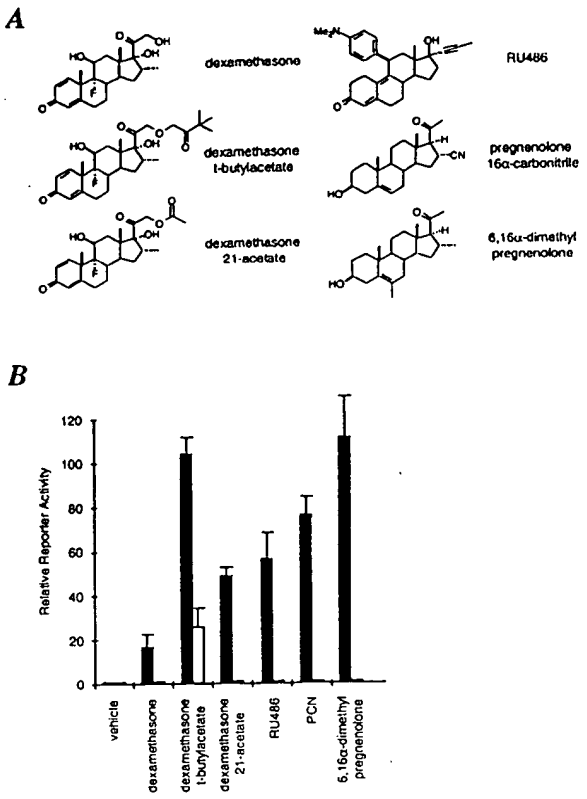


Figure 3. Synthetic Glucocorticoids and Pregnenolone Derivatives Activate PXR

(A) Structures of the steroids that activate the GAL4-PXR chimeric proteins.

(B) CV-1 cells were cotransfected with expression plasmids encoding either GAL4-PXR.1 (filled bars) or GAL4-PXR.2 (open bars) and the reporter plasmid *(UAS)$_5$-tk-CAT*. Cells were treated with vehicle alone (0.1% DMSO) or 10 μM of the indicated steroids. Cell extracts were subsequently assayed for CAT activity. Data represent the mean of five data points from two different experiments ± SD.

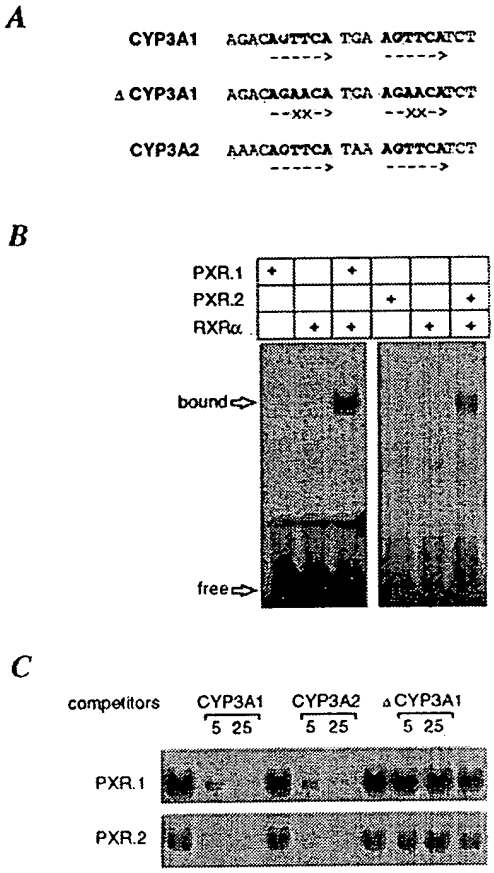


Figure 4. PXR Binds as a Heterodimer with RXRα to DR-3 Response Elements Present in the *CYP3A1* and *CYP3A2* Gene Promoters

(A) Alignment of DR-3 motifs present in the promoter regions of the *CYP3A1* and *CYP3A2* genes. A mutated derivative of the *CYP3A1* DR-3 motif (Δ*CYP3A1*) used in the gel mobility shift assays and the position of the mutations is also shown.

(B) Gel mobility shift assays were performed with a radiolabeled oligonucleotide containing the *CYP3A1* DR-3 and in vitro synthesized PXR.1, PXR.2, and RXRα as indicated.

(C) Gel mobility shift assays were performed with PXR.1 and RXRα (top panel) or PXR.2 and RXRα (bottom panel) and radiolabeled oligonucleotide containing the *CYP3A1* DR-3 in the presence of either a 5-fold or 25-fold excess of unlabeled oligonucleotides containing the *CYP3A1* DR-3, a mutated *CYP3A1* DR-3 (Δ*CYP3A1*), or the *CYP3A2* DR-3 as indicated.

of transcription and has been mapped to a conserved region within the *CYP3A1* and *CYP3A2* gene promoters that does not contain a typical glucocorticoid response element, but instead contains a DR of the nonsteroid nuclear receptor half-site sequence AGTTCA separated by a three-nucleotide spacer, a so-called DR-3 motif (Figure 4A) (Umesono et al., 1991; Miyata et al., 1995; Quattrochi et al., 1995; Huss et al., 1996). These data have led to speculation that a nonsteroid nuclear hormone receptor might be involved in mediating the effects of PCN and dexamethasone.

Given that PXR shares a high degree of homology with ONR1 and VDR in the DBD (Figure 1B) and that both ONR1 and VDR preferentially bind to DR-3 HREs as heterodimers with RXR (Umesono et al., 1991; Smith et al., 1994; Mangelsdorf and Evans, 1995), we postulated that PXR might bind to the *CYP3A1* and *CYP3A2* DR-3 motifs as a heterodimer with RXR. In order to test this idea, gel mobility shift assays were performed using a radiolabeled oligonucleotide containing the *CYP3A1* DR-3 motif and in vitro synthesized PXR.1, PXR.2, and RXRα. Neither PXR nor RXRα bound to the *CYP3A1* DR-3 alone. However, both PXR.1 and PXR.2 bound efficiently to the *CYP3A1* DR-3 as heterodimers with RXRα (Figure 4B). The PXR-RXRα complex with DNA was competed efficiently by an excess of unlabeled *CYP3A1* DR-3 oligonucleotide or an oligonucleotide containing a closely related DR-3 motif from the *CYP3A2* gene promoter (Figures 4A and 4C). No competition was seen with an oligonucleotide containing a mutated *CYP3A1* DR-3 motif (Figures 4A and 4C). Thus, both PXR isoforms can bind specifically as heterodimers with RXRα to DR-3 motifs found in the promoter regions of *CYP3A* genes.

We next asked whether the PXR isoforms could induce gene expression through the *CYP3A1* DR-3 element in response to steroids. Transient transfection assays were performed with a reporter plasmid containing two copies of the *CYP3A1* DR-3 motif inserted

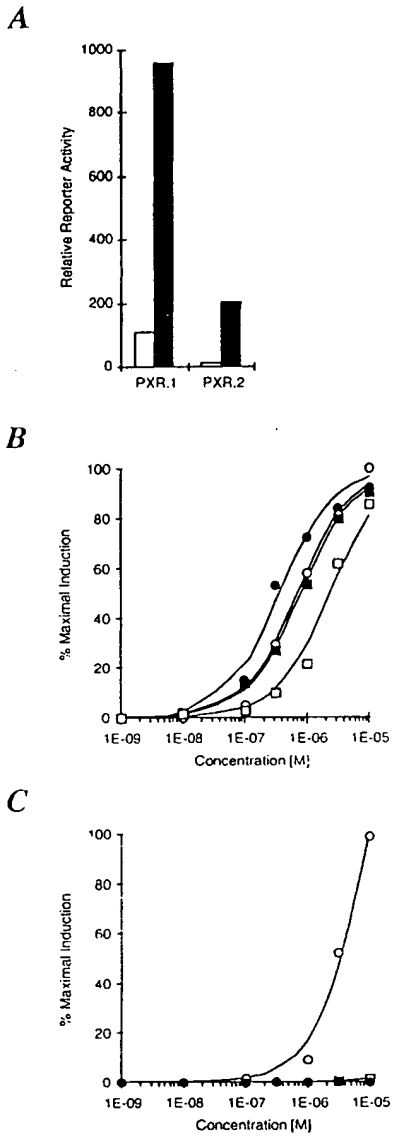


Figure 5. Synthetic Glucocorticoids and Pregnenolone Derivatives Activate PXR through the *CYP3A1* DR-3

(A) CV-1 cells were cotransfected with expression plasmids for full-length PXR.1 or PXR.2 and the *(CYP3A1)$_2$-tk-CAT* reporter plasmid. Cells were treated with either vehicle (0.1% DMSO) alone (open bars) or 10 μM dexamethasone-t-butylacetate (filled bars). Cell extracts were subsequently assayed for CAT activity.

(B and C) CV-1 cells were cotransfected with expression plasmids for full-length PXR.1 (B) or PXR.2 (C) and the *(CYP3A1)$_2$-tk-CAT* reporter plasmid. Transfected cells were treated with the indicated concentrations of 6,16α-dimethyl pregnenolone (closed circles), dexamethasone-t-butylacetate (open circles), PCN (closed squares), or RU486 (open squares). Cell extracts were subsequently assayed for CAT activity and data plotted as the percentage of the maximal response obtained. Data points represent the mean of assays performed in triplicate. Similar results were obtained in two independent experiments.

upstream of the minimal thymidine kinase promoter and the *CAT* gene [*(CYP3A1)$_2$-tk-CAT*]. Interestingly, in the absence of steroids, PXR.1 was found to have a roughly 20-fold higher basal level of activity than PXR.2 (Figure 5A). Nevertheless, both PXR isoforms were efficiently activated by dexamethasone-t-butylacetate (Figure 5A). Dose–response analysis revealed dexamethasone-t-butylacetate to be a significantly more potent activator of PXR.1 than PXR.2, with EC_{50} values of 0.8 and 5 μM for PXR.1 and PXR.2, respectively (Figures 5B and 5C). Consistent with the results obtained using the GAL4-PXR chimeras, 6,16α-dimethyl pregnenolone, PCN, and RU486 activated full-length PXR.1 on the *CYP3A1* response element but failed to activate full-length PXR.2 (Figures 5B and 5C). 6,16α-dimethyl pregnenolone was the most potent of these compounds, activating PXR.1 with an EC_{50} value of 300 nM (Figure 5B). Based upon these data, we conclude that the full-length PXR isoforms can activate gene expression through the *CYP3A1* DR-3 motif in response to synthetic steroids.

PXRs Are Activated by Naturally Occurring Steroids

As PXR was activated by synthetic steroids, we examined whether a naturally occurring steroid might serve as the endogenous hormone for PXR. Accordingly, CV-1 cells were transfected with expression plasmids for either of the full-length PXR isoforms and the *(CYP3A1)$_2$-tk-CAT* reporter and treated with a variety of natural steroids, including progestins, glucocorticoids, mineralocorticoids, androgens, estrogens, bile acids, and oxysterols. Both PXR.1 and PXR.2 were activated by micromolar concentrations of certain of these steroids. Consistent with our findings that the pregnenolone derivatives 6,16α-dimethyl pregnenolone and PCN activate PXR.1, pregnenolone and its metabolites 17α-hydroxypregnenolone, progesterone, 17α-hydroxyprogesterone, and 5β-pregnane-3,20-dione activated PXR.1 (Figure 6B). Dose–response analysis revealed that all five of these pregnanes activated PXR.1 with EC_{50} values in the 5–20 μM range (Figure 6C). Since PXR.1 is activated by synthetic glucocorticoids, we were surprised to find that naturally occurring glucocorticoids, including cortisol and corticosterone, had virtually no effect on PXR.1 activity (data not shown).

In analogous cotransfection experiments performed with PXR.2, only 5β-pregnane-3,20-dione was found to induce reporter activity >5-fold (Figure 6B). Notably, pregnenolone, progesterone, and their 17α-hydroxylated derivatives, which were efficacious activators of PXR.1, had little or no activity on PXR.2 (Figure 6B). Thus, while both PXR isoforms are activated by naturally occurring pregnanes, PXR.1 is activated by a broader range of these steroids than PXR.2. The generation of LBD isoforms with distinct activation profiles provides a novel mechanism for increasing the regulatory diversity in the PXR signaling pathway.

Steroids Promote the Interaction of PXR with a Coactivator Protein

We next sought to address whether the steroids that activated PXR did so through direct interactions with the LBD. Due to the lack of radiolabeled derivatives of the more potent PXR activators, we were unable to perform standard binding analyses. However, a number of laboratories have recently demonstrated that ligands

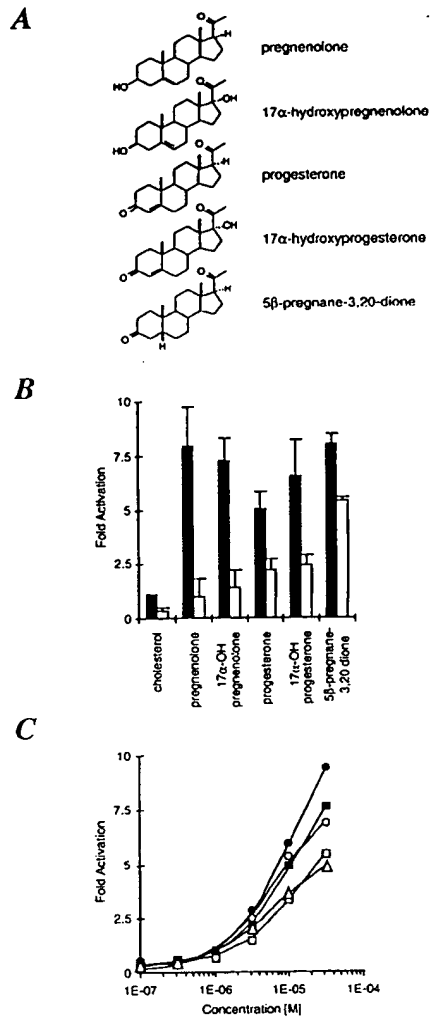


Figure 6. Naturally Occurring Steroids Activate PXR

(A) Structures of naturally occurring steroids that activate PXR.

(B) CV-1 cells were cotransfected with expression plasmids for full-length PXR.1 (filled bars) or PXR.2 (open bars) and the *(CYP3A1)$_2$-tk-CAT* reporter plasmid. Transfected cells were treated with 10 μM of the indicated steroids. Cell extracts were subsequently assayed for CAT activity and data plotted as fold-activation relative to cells treated with vehicle (0.1% DMSO) alone. Data points represent the mean of assays performed in triplicate ± SD.

(C) Transfected cells were treated with the indicated concentrations of pregnenolone (closed circles), 17α-hydroxypregnenolone (open circles), progesterone (triangles), 17α-hydroxyprogesterone (open squares), or 5β-pregnane-3,20-dione (closed squares). Cell extracts were subsequently assayed for CAT activity and data plotted as fold-activation relative to cells treated with vehicle (0.1% DMSO) alone. Data points represent the mean of assays performed in triplicate. Similar results were obtained in two independent experiments.

for nuclear receptors induce their interaction with proteins required for optimal transcriptional activation, the so-called coactivator proteins (Horwitz et al., 1996). These ligand-dependent interactions have been exploited as a biochemical assay for demonstrating direct interactions between ligands and their cognate receptors (Krey et al., 1997).

The steroid receptor coactivator protein-1 (SRC-1) has been shown to interact directly with both steroid and nonsteroid nuclear hormone receptors in a ligand-dependent fashion (Onate et al., 1995; Horwitz et al., 1996; Kamei et al., 1996; Takeshita et al., 1996). The interaction of SRC-1 with nuclear receptors is dependent upon the amino acid motif LXXLL found in multiple copies in SRC-1 (Heery et al., 1997; Torchia et al., 1997). In an effort to determine whether the steroid activators of PXR serve as ligands for this orphan receptor, we tested whether a 14 kDa fragment of SRC-1 (SRC-1.14) containing three LXXLL motifs interacted with PXR in a steroid-dependent manner.

SRC-1.14 was expressed in vitro and labeled with [^{35}S]-methionine and [^{35}S]-cysteine, and the LBD of PXR.1 was expressed in *E. coli* as a fusion protein with glutathione-S-transferase (GST). Coprecipitation experiments were performed in the presence of the most potent PXR activators including 6,16α-dimethyl pregnenolone, dexamethasone-t-butylacetate, and PCN. [^{35}S]-SRC-1.14 interacted only weakly with the GST-PXR.1LBD fusion protein in the absence of added compound (Figure 7A). The interaction of [^{35}S]-SRC-1.14 with GST-PXR.1LBD was significantly enhanced in the presence of either dexamethasone-t-butylacetate, 6,16α-dimethyl pregnenolone, or PCN (Figure 7A). Additional experiments performed with 6,16α-dimethyl pregnenolone and PCN revealed that these steroids promoted [^{35}S]-SRC-1.14/PXR.1LBD interactions in a dose-dependent manner (Figure 7B). Consistent with the results of the transfection studies, 6,16α-dimethyl pregnenolone was more potent than PCN in promoting these interactions (Figure 7B). Estradiol did not stimulate interactions between PXR.1 and [^{35}S]-SRC-1.14, indicating that the enhancement was specific for compounds that activate PXR.1 in the transfection assay (Figure 7A). In control experiments, estradiol promoted the efficient interaction of [^{35}S]-SRC-1.14 with a GST-ERαLBD fusion protein (Figure 7A). However, GST-ERαLBD interactions with [^{35}S]-SRC-1.14 were not induced in the presence of PCN, dexamethasone-t-butylacetate, or 6,16α-dimethyl pregnenolone. Taken together, these data provide strong evidence that PCN, dexamethasone-t-butylacetate, and 6,16α-dimethyl pregnenolone serve as ligands for PXR.

Discussion

Identification of a Novel Signaling Pathway for Steroids

It was first observed over 50 years ago that repeated administration of high doses of certain steroids, including progestins and androgens, reduced their own and each others toxic effects. These early observations, together with the later findings that resistance to numerous drugs is sex dependent, that castration of rodents leads to increased drug sensitivity, and that liver homogenates of intact male rats metabolize certain drugs more rapidly than homogenates prepared from castrated animals led to the concept of "catatoxic" steroids; that is, steroids that confer resistance to specific toxins by accelerating their metabolism (reviewed by Kourounakis et al., 1977). It was speculated that catatoxic agents might have utility in the treatment of patients suffering from either drug intoxication or from diseases caused by endogenous substances liable to metabolism (e.g., Cushing's syndrome) (Kourounakis et al., 1977).

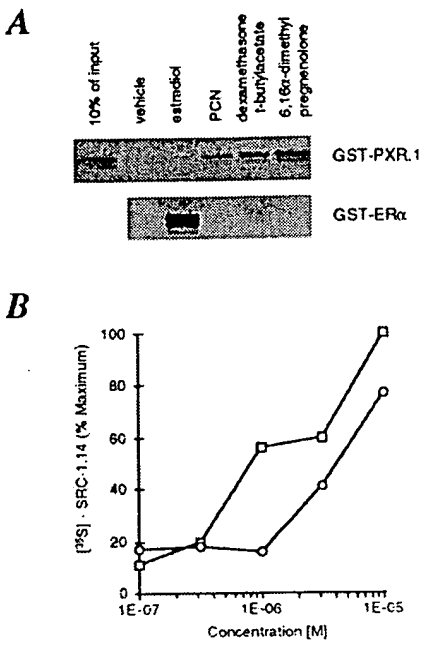


Figure 7. Steroids Induce PXR Interactions with a Fragment of the Coactivator Protein SRC-1

(A) Coprecipitation experiments were performed with bacterially expressed GST-PXR.1LBD (upper panel) or GST-ERαLBD (lower panel) and [^{35}S]-SRC-1.14 synthesized in vitro. [^{35}S]-SRC-1.14 was mixed with either GST-PXR.1LBD or GST-ERαLBD in the presence of vehicle alone (1% DMSO) or 10 μM of estradiol, PCN, dexamethasone-t-butylacetate, or 6,16α-dimethyl pregnenolone as indicated. [^{35}S]-SRC-1.14 complexed with either GST-PXR.1LBD or GST-ERαLBD was precipitated with glutathione-sepharose beads as described in Experimental Procedures. A lane representing 10% of the input [^{35}S]-SRC-1.14 in each reaction is shown on the left. Western blot analysis with an anti-GST antibody revealed that comparable amounts of GST-PXR.1LBD and GST-ERαLBD fusion proteins were used in the assays (data not shown).

(B) Dose–response analysis was performed with [^{35}S]-SRC-1.14 and GST-PXR.1LBD in the presence of the indicated concentrations of 6,16α-dimethyl pregnenolone (squares) or PCN (circles). [^{35}S]-SRC-1.14 was quantitated via scanning densitometry and plotted as a percent of the [^{35}S]-SRC-1.14 precipitated in the presence of 10 μM 6,16α-dimethyl pregnenolone. Data shown represent the average of duplicate points, and similar results were obtained in two separate experiments.

A systematic analysis of steroids in the early 1970s identified the pregnenolone derivative PCN as the most potent catatoxic compound among those tested (Selye, 1971). Insight into the mechanism underlying the catatoxic effects of PCN was provided by the demonstration that this synthetic steroid induces the expression of CYP3A1 and CYP3A2, two closely related members of the P450 family of monooxygenases (Elshourbagy and Guzelian, 1980; Heuman et al., 1982; Hardwick et al., 1983; Schuetz and Guzelian, 1984; Schuetz et al., 1984; Gonzalez et al., 1985). The CYP3A hemoproteins have a remarkably broad substrate specificity, hydroxylating a variety of xenobiotics such as cyclosporin, warfarin, and erythromycin, as well as endogenous steroids including cortisol, progesterone, testosterone, and DHEA-sulfate (Nebert and Gonzalez, 1987; Juchau, 1990). Subsequent studies with the cloned *CYP3A1* gene promoter identified a PCN response element that was highly conserved in the *CYP3A2* gene promoter (Miyata et al., 1995; Quattrochi et al., 1995). This response element was composed of two copies of the nuclear receptor half-site consensus sequence AGTTCA organized as a DR.

In addition to PCN, the expression of *CYP3A1* was also shown to be induced by dexamethasone both in vivo and in cultured hepatocytes (Heuman et al., 1982; Schuetz and Guzelian, 1984; Schuetz et al., 1984). However, the concentrations of dexamethasone required to induce *CYP3A* gene expression were higher than those typically required to activate the classical GR signaling pathway (Schuetz and Guzelian, 1984; Schuetz et al., 1984). Promoter mapping studies showed that dexamethasone induced *CYP3A1* gene expression through the same DR response element as PCN (Quattrochi et al., 1995; Huss et al., 1996). Thus, paradoxically, both high concentrations of dexamethasone, a glucocorticoid, and PCN, an antiglucocorticoid, induced the expression of the *CYP3A1* gene through the same response element.

We now provide several lines of evidence indicating that the orphan nuclear receptor PXR is responsible for mediating the inductive effects of PCN and dexamethasone on *CYP3A* gene expression. First, both dexamethasone and PCN are efficacious activators of the PXR.1 isoform. Second, PXR binds efficiently as a heterodimer with RXR to the conserved DR-3 motifs identified in the *CYP3A1* and *CYP3A2* gene promoters as glucocorticoid and PCN response elements. Finally, we detected PXR expression in only a few tissues, including liver, intestine, and kidney. These are the primary tissues in which the *CYP3A* genes are expressed and induced in response to either dexamethasone or PCN treatment (Heuman et al., 1982; Debri et al., 1995). Our data thus support the existence of a novel signaling pathway for synthetic glucocorticoids and provide a mechanistic explanation for the long-standing paradox as to how both GR agonists and antagonists can exert similar effects on *CYP3A* gene expression. Moreover, the identification of PXR.1 as the PCN receptor provides the tool necessary for the rapid identification of novel pharmacological agents with more potent catatoxic activities.

In addition to inducing *CYP3A* gene expression, PCN is also known to have marked effects on hepatic cholesterol homeostasis in rodents. These effects include significant decreases in the levels of HMG-CoA reductase and cholesterol 7α-hydroxylase gene expression with concomitant reductions in sterol biosynthesis and bile acid secretion (von Bergmann et al., 1975; Mason and Boyd, 1978; Turley and Dietschy, 1984; Stahlberg, 1995). PCN has also been reported to enhance the formation of cholesterol-esters and the hypersecretion of cholesterol into the bile (von Bergmann et al., 1975; Turley and Dietschy, 1984). Thus, PCN affects key aspects of cholesterol metabolism including its biosynthesis, storage, and secretion. Although we cannot exclude the possibility that some of its biological effects might be mediated through the GR or other steroid receptors, it is tempting to speculate that PCN is mimicking the actions of an endogenous hormone that serves to regulate coordinately steroid and sterol metabolism through the activation of PXR in tissues such as liver and intestine. Our data raise the possibility of the existence of regulatory loops through which endogenous PXR hormones feed back to regulate cholesterol homeostasis and feed forward to regulate steroid homeostasis.

What is the natural ligand for PXR? Pregnenolone and progesterone are among the most potent naturally occurring PXR activators that we have identified, activating PXR.1 at low micromolar concentrations. Pregnenolone is one of the most abundant steroids in mammals, circulating in humans and rodents at concentrations that range from roughly 1 to 50 nM (Punjabi et al., 1983; Wichmann et al., 1984; Tietz, 1995). While progesterone levels in human serum are generally in the 1–10 nM range, concentrations can exceed 700 nM during the third trimester of pregnancy (Tietz, 1995). These levels approach those required to activate PXR.1 in vitro. Nevertheless, it remains unclear whether concentrations of progesterone or pregnenolone sufficient to activate PXR.1 are achieved either in serum or in tissues under normal physiological conditions. Our results with PXR may thus be analogous to those obtained with RXR, which was first shown to be activated by micromolar concentrations of all-*trans* retinoic acid (t-RA) prior to the identification of 9-*cis* RA as a high-affinity ligand (Mangelsdorf et al., 1990; Heyman et al., 1992; Levin et al., 1992). Based upon our findings that PXR is activated by pregnenolone and its metabolites and that the synthetic steroid 6,16α-dimethyl pregnenolone activates PXR with an EC_{50} value of 300 nM, we suggest that the natural PXR ligand is likely to be a pregnane.

Perspectives

With the isolation of the androgen receptor in 1988 (Chang et al., 1988; Lubahn et al., 1988), receptors for all of the known nuclear-acting steroid hormones had been cloned. However, studies performed during the past two years with orphan members of the nuclear receptor family have suggested that additional sterols are likely to serve as mammalian hormones. For example, the orphan receptors LXR and SF-1 were recently shown to be activated by physiological concentrations of several oxysterol metabolites of cholesterol (Janowski et al., 1996; Lala et al., 1997; Lehmann et al., 1997). While the biological role of LXR remains less clear, SF-1 is essential for adrenal and gonadal development and regulates the expression of genes required for steroidogenesis (Luo et al., 1994). These data suggested an unexpected hormonal function for oxysterols in the regulation of steroidogenesis.

We have now identified PXR as a novel member of the nuclear receptor family that is efficaciously activated by both natural and synthetic steroids. The activation profile of PXR is distinct from any of the other steroid hormone receptors identified to date, suggesting that this orphan receptor defines a novel endocrine signaling pathway. We conclude that the identification of PXR provides additional evidence for an expanded role for steroid hormones in mammalian physiology and that the elucidation of the biological role of PXR is likely to lead to a better understanding of how steroids elicit their myriad effects.

Experimental Procedures

Chemicals

Dexamethasone-t-butylacetate and RU486 were purchased from Research Plus, Inc. (Bayonne, NJ) and Biomol (Plymouth Meeting, PA), respectively. All other steroids were purchased from either Sigma Chemical Co. (St. Louis, MO) or Steraloids, Inc. (Wilton, NH).

Molecular Cloning of *PXR* cDNAs

Partial mouse sequence for PXR was obtained from an EST deposited in the Washington University/HHMI EST database (accession number AA277370). A 19 bp oligonucleotide derived from this EST sequence (5′ TCTCCCCAGATCGTCCTGG 3′) was used to screen a pCMV-SPORT mouse liver cDNA library (GIBCO-BRL) using Gene Trapper solution hybridization cloning technology (GIBCO-BRL). Five clones were obtained that ranged in size from 1.0 kb to 2.7 kb. Four of these clones encoded PXR.1 and one encoded the DBD and LBD of PXR.2. A clone encoding the full-length PXR.2 isoform-coding region (nucleotides 151 to 1443) was subsequently isolated from mouse liver cDNA through PCR. Sequences were aligned and analyzed by the University of Wisconsin Genetics Computer Group programs.

Plasmids

The expression plasmids pSG5-GAL4-PXR.1LBD and pSG5-GAL4-PXR.2LBD were generated by amplification of cDNA encoding amino acids 105–431 of PXR.1 or 105–390 of PXR.2 by PCR and insertion into a modified pSG5 expression vector (Stratagene) containing DNA encoding the DBD of GAL4 (amino acids 1–147) and the SV40 Tag nuclear localization signal (APKKKRKVG) inserted upstream of a multiple cloning site. The *$(UAS)_5$-tk-CAT* reporter plasmid has been previously described (Lehmann et al., 1995). The expression vectors pSG5-PXR.1 and pSG5-PXR.2 were generated by amplification of cDNA encoding amino acids 1–431 of PXR.1 or 1–390 of PXR.2 and insertion into pSG5. The reporter plasmid *$(CYP3A1)_2$-tk-CAT* was generated by insertion of two copies of a double-stranded oligonucleotide containing the *CYP3A1* DR-3 RE (5′ GATCAGACAGTTCATGAAGTTCATCTAGATC 3′) (Quattrochi et al., 1995; Huss et al., 1996) into the BamHI site of pBLCAT2 (Luckow and Schütz, 1987). The bacterial expression vector pGEX-PXR.1LBD was generated by PCR amplification of cDNA encoding amino acids 105–431 of PXR.1 and insertion into a pGEX-2T vector (Pharmacia) that had been modified to contain KpnI and NotI restriction sites. The bacterial expression vector pGEX-ERαLBD was generated by PCR amplification of cDNA encoding amino acids 251–595 of human ERα (Green et al., 1986) and insertion into pGEX-2T. The expression plasmid for the SRC-1.14 fragment was generated by PCR amplification of DNA encoding amino acids 632–754 of human SRC-1 (Takeshita et al., 1996) and insertion into the expression vector pRSETC (Invitrogen). All constructs were confirmed by sequence analysis.

Cotransfection Assays

CV-1 cells were plated in 24-well plates in DME medium supplemented with 10% charcoal-stripped fetal calf serum at a density of 1.2×10^5 cells per well. In general, transfection mixes contained 33 ng of receptor expression vector, 100 ng of reporter plasmid, 200 ng of β-galactosidase expression vector (pCH110, Pharmacia), and 166 ng of carrier plasmid. Cells were transfected overnight by lipofection using Lipofectamine (Life Technologies, Inc.) according to the manufacturer's instructions. The medium was changed to DME medium supplemented with 10% delipidated calf serum (Sigma) and cells were incubated for an additional 24 hr. Cell extracts were prepared and assayed for CAT and β-galactosidase activities as described previously (Lehmann et al., 1995).

Northern Analysis

An approximately 1.0 kb fragment encoding the LBD of PXR.1 (nucleotides 463–1446) was ^{32}P-labeled by random priming using the T7 Quick-Prime kit (Pharmacia). This radiolabeled fragment was used to probe mouse multiple-tissue Northern blots (OriGene, Rockville, MD). Blots were hybridized in ExpressHyb solution (Clontech, Palo Alto, CA) at 65°C for 18 hr. Final washes were performed with 0.1× SSC, 0.1% SDS at 65°C.

In Situ Hybridization Analysis

Embryonic day 18 (E18) mice (C57bl/CBA and NMRI, Bomholt Breeding and Research Center, Copenhagen, Denmark) were used in the studies. Tissues were sectioned at 14 μm and thaw-mounted

onto slides (ProbeOn, Fischer). Three different oligonucleotide probes (Pharmacia Biotech, Sollentuna, Sweden) designed to hybridize to *PXR* mRNA were used. Oligonucleotides were radiolabeled using [^{35}S]-deoxyadenosine 5'-α-thio-triphosphate (NEN) at the 3' end using terminal deoxynucleotidyl transferase (Amersham). Labeled oligonucleotides were hybridized to tissue sections and mRNA expression was first detected by phosphoimaging (BASF 3000 Phosphoimager, Fuji) followed by film emulsion autoradiography (Dagerlind et al., 1992). Sections were examined using light- and dark-field microscopy (Axiophot, Zeiss) and photographed (Ektachrome 64T, Kodak). Positives were scanned (SprintScan 35, Polaroid) and processed using Photoshop and PageMaker, Adobe. The following oligonucleotide sequences were used as probes: probe 1 (PXR.1 and PXR.2 specific), 5' GGAGCCTCAATCTTTTCCC TCTTCTTCCTCTTGATCAAGGCCCGC 3'; probe 2 (PXR.1 specific), 5' CTTCCAACAGTGAGGCCTGCAGAAACTCTGGAAGCTCACAGC CAC 3'; probe 3 (PXR.1 and PXR.2 specific), 5' TGGGCTCTTCCAAG GCAGAGTCTGCTTCT TCACACTGTACAAGGCC 3'. A 50 bp random oligonucleotide was used as a negative control.

Gel Mobility Shift Assays

PXR.1, PXR.2, and RXRα were synthesized in vitro using the TNT rabbit reticulocyte lysate coupled in vitro transcription/translation system (Promega) according to the manufacturer's instructions. Gel mobility shift assays (20 μl) contained 10 mM Tris (pH 8.0), 40 mM KCl, 0.05% NP-40, 6% glycerol, 1 mM DTT, 0.2 μg of poly(dI-dC), and 2.5 μl each of in vitro-synthesized PXR and RXR proteins. The total amount of reticulocyte lysate was maintained constant in each reaction (5 μl) through the addition of unprogrammed lysate. Competitor oligonucleotides were included at a 5-fold or 25-fold excess as indicated in the figure legends. After a 10 min incubation on ice, 10 ng of ^{32}P-labeled oligonucleotide was added and the incubation continued for an additional 10 min. DNA–protein complexes were resolved on a 4% polyacrylamide gel in 0.5× TBE (1× TBE = 90 mM Tris, 90 mM boric acid, 2 mM EDTA). Gels were dried and subjected to autoradiography at -70°C. The following double-stranded oligonucleotides were used as either radiolabeled probes or competitors: *CYP3A1*, 5' GATGCAGACAGTTCATGAAGTTCATCT AGATC 3' (Quattrochi et al., 1995); *ΔCYP3A1*, 5' GATCAGACAGAAC ATGAAGAACATCTAGATC 3'; *CYP3A2*, 5' GATCAAACAGTTCATAA AGTTCATCTAGATC 3' (Miyata et al., 1995).

SRC-1.14 Coprecipitation Assay

GST-PXR.1LBD and GST-ERαLBD fusion proteins were expressed in BL21(DE3)plysS cells and bacterial extracts prepared by one cycle of freeze-thaw of the cells in protein lysis buffer containing 10 mM Tris (pH 8.0), 50 mM KCl, 10 mM DTT, and 1% NP-40 followed by centrifugation at 40,000 × g for 30 min. Glycerol was added to the resulting supernatant to a final concentration of 10%. Lysates were stored at −80°C. [^{35}S]-SRC-1.14 was generated using the TNT rabbit reticulocyte system (Promega) in the presence of Pro-Mix (Amersham). Coprecipitation reactions included 25 μl of lysate containing GST-PXR.1LBD or GST-ERαLBD fusion proteins or control GST; 25 μl incubation buffer (50 mM KCl, 40 mM HEPES [pH 7.5]; 5 mM β-mercaptoethanol; 1% Tween-20, 1% non-fat dry milk); 5 μl [^{35}S]-SRC-1.14; and either PCN; dexamethasone-t-butylacetate; 6,16α-dimethyl pregnenolone; estradiol; or control DMSO. The mixtures were incubated for 25 min at 4°C with gentle mixing prior to the addition of 15 μl of glutathione-sepharose 4B beads (Pharmacia) that had been extensively washed with protein lysis buffer. Reactions were incubated with gentle mixing at 4°C for an additional 20 min. The beads were pelleted at 2000 rpm in a microfuge and washed 3 times with protein lysis buffer containing either PCN; dexamethasone-t-butylacetate; 6,16α-dimethyl pregnenolone; estradiol; or control DMSO. After the last wash, the beads were resuspended in 25 μl of 2× SDS-PAGE sample buffer containing 1 mM DTT. Samples were heated at 100°C for 5 min and loaded onto a 10% Bis-Tris PAGE gel. Autoradiography was performed for 2–4 days. Assays were quantitated using a Molecular Dynamics Computing Densitometer and Image Quant software.

Acknowledgments

We thank Eva Lindqvist for expert technical assistance with the in situ hybridization studies; Kelli Plunket for expert assistance with the transient transfection studies and data analysis; Nick Tomkinson for assistance in compound selection; Dave Morris and Lars Olson for support and encouragement throughout these studies; and John Didsbury for comments on the manuscript. This work was supported by grants from the Swedish Research Council to T. P. and R. H. Z.

Received October 30, 1997; revised November 20, 1997.

References

Baker, A.R., McDonnell, D.P., Hughes, M., Crisp, T.M., Mangelsdorf, D.J., Haussler, M.R., Pike, J.W., Shine, J., and O'Malley, B.W. (1988). Cloning and expression of full-length cDNA encoding human vitamin D receptor. Proc. Natl. Acad. Sci. USA *85*, 3294–3298.

Beato, M., Herrlich, P., and Schütz, G. (1995). Steroid hormone receptors: many actors in search of a plot. Cell *83*, 851–857.

Cadepond, F., Ulmann, A., and Baulieu, E.E. (1997). RU486 (mifepristone): mechanisms of action and clinical uses. Annu. Rev. Med. *48*, 129–156.

Chang, C.S., Kokontis, J., and Liao, S.T. (1988). Molecular cloning of human and rat complementary DNA encoding androgen receptors. Science *240*, 324–326.

Dagerlind, A., Friberg, K., Bean, A.J., and Hokfelt, T. (1992). Sensitive mRNA detection using unfixed tissue: combined radioactive and non-radioactive in situ hybridization histochemistry. Histochemistry *98*, 39–49.

Debri, K., Boobis, A.R., Davies, D.S., and Edwards, R.J. (1995). Distribution and induction of CYP3A1 and CYP3A2 in rat liver and extrahepatic tissues. Biochem. Pharmacol. *50*, 2047–2056.

Elshourbagy, N.A., and Guzelian, P.S. (1980). Separation, purification, and characterization of a novel form of cytochrome P-450 from rats treated with pregnenolone-16α-carbonitrile. J. Biol. Chem. *255*, 1279–1285.

Enmark, E., and Gustafsson, J.-A. (1996). Orphan nuclear receptors—the first eight years. Mol. Endocrinol. *10*, 1293–1307.

Evans, R.M. (1988). The steroid and thyroid hormone receptor superfamily. Science *240*, 889–895.

Forman, B.M., Goode, E., Chen, J., Oro, A.E., Bradley, D.J., Perlmann, T., Noonan, D.J., Burka, L.T., McMorris, T., Lamph, W.W., et al. (1995). Identification of a nuclear receptor that is activated by farnesol metabolites. Cell *81*, 687–693.

Forman, B.M., Chen, J., and Evans, R.M. (1997). Hypolipidemic drugs, polyunsaturated fatty acids, and eicosanoids are ligands for peroxisome proliferator-activated receptors α and δ. Proc. Natl. Acad. Sci. USA *94*, 4312–4317.

Glass, C.K. (1994). Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. Endocr. Rev. *15*, 391–407.

Gonzalez, F.J., Nebert, D.W., Hardwick, J.P., and Kasper, C.B. (1985). Complete cDNA and protein sequence of a pregnenolone 16 alpha-carbonitrile-induced cytochrome P-450. A representative of a new gene family. J. Biol. Chem. *260*, 7435–7441.

Green, S., Walter, P., Kumar, V., Krust, A., Bornet, J.-M., Argos, P., and Chambon, P. (1986). Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. Nature *320*, 134–139.

Hardwick, J.P., Gonzalez, F.J., and Kasper, C.B. (1983). Cloning of DNA complementary to cytochrome P-450 induced by pregnenolone-16 alpha-carbonitrile. Characterization of its mRNA, gene, and induction response. J. Biol. Chem. *258*, 10182–10186.

Heery, D.M., Kalkhoven, E., Hoare, S., and Parker, M.G. (1997). A signature motif in transcriptional co-activators mediates binding to nuclear receptors. Nature *387*, 733–736.

Heuman, D.M., Gallagher, E.J., Barwick, J.L., Elshourbagy, N.A., and Guzelian, P.S. (1982). Immunochemical evidence for induction of a common form of hepatic cytochrome P-450 in rats treated with pregnenolone-16α-carbonitrile or other steroidal or non-steroidal agents. Mol. Pharmacol. *21*, 753–760.

Heyman, R.A., Mangelsdorf, D.J., Dyck, J.A., Stein, R., Eichele, G.,

Evans, R.M., and Thaller, C. (1992). 9-*cis* retinoic acid is a high affinity ligand for the retinoid X receptor. Cell *68*, 397–406.

Horwitz, K.B., Jackson, T.A., Bain, D.L., Richer, J.K., Takimoto, G.S., and Tung, L. (1996). Nuclear receptor coactivators and corepressors. Mol. Endocrinol. *10*, 1167–1177.

Huss, J.M., Wang, S.I., Astrom, A., McQuiddy, P., and Kasper, C.B. (1996). Dexamethasone responsiveness of a major glucocorticoid-inducible CYP3A gene is mediated by elements unrelated to a glucocorticoid receptor binding motif. J. Biol. Chem. *93*, 4666–4670.

Janowski, B.A., Willy, P.J., Devi, T.R., Falck, J.R., and Mangelsdorf, D.J. (1996). An oxysterol signaling pathway mediated by the nuclear receptor LXRα. Nature *383*, 728–731.

Juchau, M.R. (1990). Substrate specificities and functions of the P450 cytochromes. Life Sci. *47*, 2385–2394.

Kamei, Y., Xu, L., Heinzel, T., Torchia, J., Kurokawa, R., Gloss, B., Lin, S.-C., Heyman, R.A., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (1996). A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. Cell *85*, 403–414.

Kliewer, S.A., Sundseth, S.S., Jones, S.A., Brown, P.J., Wisely, G.B., Koble, C.S., Devchand, P., Wahli, W., Willson, T.M., Lenhard, J.M., and Lehmann, J.M. (1997). Fatty acids and eicosanoids regulate gene expression through direct interactions with peroxisome proliferator-activated receptors α and γ. Proc. Natl. Acad. Sci. USA *94*, 4318–4323.

Kourounakis, P., Selye, H., and Tache, Y. (1977). Catatoxic steroids. Adv. Steroid Biochem. Pharmacol. *6*, 35–57.

Krey, G., Braissant, O., L'Horset, F., Kalkhoven, E., Perroud, M., Parker, M.G., and Wahli, W. (1997). Fatty acids, eicosanoids, and hypolipidemic agents identified as ligands of peroxisome proliferator-activated receptors by coactivator-dependent receptor ligand assay. Mol. Endocrinol. *11*, 779–791.

Lala, D.S., Syka, P.M., Lazarchik, S.B., Mangelsdorf, D.J., Parker, K.L., and Heyman, R.A. (1997). Activation of the nuclear orphan receptor steroidogenic factor 1 by oxysterols. Proc. Natl. Acad. Sci. USA *94*, 4895–4900.

Lehmann, J.M., Moore, L.B., Smith-Oliver, T.A., Wilkison, W.O., Willson, T.M., and Kliewer, S.A. (1995). An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor γ (PPARγ). J. Biol. Chem. *270*, 12953–12956.

Lehmann, J.M., Kliewer, S.A., Moore, L.B., Smith-Oliver, T.A., Oliver, B.B., Su, J.-L., Sundseth, S.S., Winegar, D.A., Blanchard, D.E., Spencer, T.A., and Willson, T.M. (1997). Activation of the nuclear receptor LXR by oxysterols defines a new hormone response pathway. J. Biol. Chem. *272*, 3137–3140.

Levin, A.A., Sturzenbecker, L.J., Kazmer, S., Bosakowski, T., Huselton, C., Allenby, G., Speck, J., Kratzeisen, C., Rosenberger, M., Lovey, A., and Grippo, J.F. (1992). 9-Cis retinoic acid stereoisomer binds and activates the nuclear receptor RXRα. Nature *355*, 359–361.

Lubahn, D.B., Joseph, D.R., Sullivan, P.M., Willard, H.F., French, F.S., and Wilson, E.M. (1988). Cloning of human androgen receptor complementary DNA and localization to the X chromosome. Science *240*, 327–330.

Luckow, B., and Schütz, G. (1987). CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. Nucleic Acids Res. *15*, 5490.

Luo, X., Ikeda, Y., and Parker, K.L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. Cell *77*, 481–490.

Mangelsdorf, D.J., and Evans, R.M. (1995). The RXR heterodimers and orphan receptors. Cell *83*, 841–850.

Mangelsdorf, D.J., Ong, E.S., Dyck, J.A., and Evans, R.M. (1990). Nuclear receptor that identifies a novel retinoic acid response pathway. Nature *345*, 224–229.

Mangelsdorf, D.J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Kastner, P., Mark, M., Chambon, P., and Evans, R.M. (1995). The nuclear receptor superfamily: the second decade. Cell *83*, 835–839.

Mason, J.I., and Boyd, G.S. (1978). The suppressive effect of the catatoxic steroid, pregnenolone-16alpha-carbonitrile, on liver microsomal cholesterol-7alpha-hydroxylase. Steroids *31*, 849–854.

Miyata, M., Nagata, K., Yamazoe, Y., and Kato, R. (1995). Transcriptional elements directing a liver-specific expression of P450/6βA (CYP3A2) gene-encoding testosterone 6β-hydroxylase. Arch. Biochem. Biophysics *318*, 71–79.

Nebert, D.W., and Gonzalez, F.J. (1987). P450 genes: structure, evolution, and regulation. Annu. Rev. Biochem. *56*, 945–993.

Onate, S.A., Tsai, S.Y., Tsai, M.-J., and O'Malley, B.M. (1995). Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. Science *270*, 1354–1357.

Pratt, W.B. (1993). The role of heat shock proteins in regulating the function, folding, and trafficking of the glucocorticoid receptor. J. Biol. Chem. *268*, 21455–21458.

Punjabi, U., Deslypere, J.P., Verdonck, L., and Vermeulen, A. (1983). Androgen and precursor levels in serum and testes of adult rats under basal conditions and after hCG stimulation. J. Steroid Biochem. *19*, 1481–1490.

Quattrochi, L.C., Mills, A.S., Barwick, J.L., Yockey, C.B., and Guzelian, P.S. (1995). A novel cis-acting element in a liver cytochrome P450 3A gene confers synergistic induction by glucocorticoids plus antiglucocorticoids. J. Biol. Chem. *270*, 28917–28923.

Schuetz, E.G., and Guzelian, P.S. (1984). Induction of cytochrome P-450 by glucocorticoids in rat liver. J. Biol. Chem. *259*, 2007–2012.

Schuetz, E.G., Wrighton, S.A., Barwick, J.L., and Guzelian, P.S. (1984). Induction of cytochrome P-450 by glucocorticoids in rat liver. J. Biol. Chem. *259*, 1999–2006.

Selye, H. (1971). Hormones and resistance. J. Pharm. Sci. *60*, 1–28.

Smith, D.P., Mason, C.S., Jones, E.A., and Old, R.W. (1994). A novel nuclear receptor subfamily member in *Xenopus* that associates with RXR and shares extensive sequence similarity to the mammalian vitamin D3 receptor. Nucleic Acids Res. *22*, 66–71.

Stahlberg, D. (1995). Effects of pregnenolone-16α-carbonitrile on the metabolism of cholesterol in rat liver microsomes. Lipids *30*, 361–364.

Takeshita, A., Yen, P.M., Misiti, S., Cardona, G.R., Liu, Y., and Chin, W.W. (1996). Molecular cloning and properties of a full-length putative thyroid hormone receptor coactivator. Endocrinology *137*, 3594–3597.

Tietz, N.W. (1995). Clinical Guide to Laboratory Tests (Philadelphia: W. B. Saunders Co.).

Torchia, J., Rose, D.W., Inostroza, J., Kamei, Y., Westin, S., Glass, C.K., and Rosenfeld, M.G. (1997). The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. Nature *387*, 677–684.

Turley, S.D., and Dietschy, J.M. (1984). Modulation of the stimulatory effect of pregnenolone-16α-carbonitrile on biliary cholesterol output in the rat by the manipulation of the rate of hepatic cholesterol synthesis. Gastroenterology *87*, 284–292.

Umesono, K., Murakami, K.K., Thompson, C.C., and Evans, R.M. (1991). Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. Cell *65*, 1255–1266.

von Bergmann, K., Schwartz, H.P., and Paumgartner, G. (1975). Effect of phenobarbital, spironolactone and pregnenolone-16α-carbonitrile on bile formation in the rat. Naunyn Schmiedebergs Arch. Pharmacol. *287*, 33–45.

Wichmann, U., Wichmann, G., and Krause, W. (1984). Serum levels of testosterone precursors, testosterone and estradiol in 10 animal species. Exp. Clin. Endocrinol. *83*, 283–290.

Wurtz, J.-M., Bourguet, W., Renaud, J.-P., Vivat, V., Chambon, P., Moras, D., and Gronemeyer, H. (1996). A canonical structure for the ligand-binding domain of nuclear receptors. Nature Struct. Biol. *3*, 87–94.

Zavacki, A.M., Lehmann, J.M., Seol, W., Willson, T.M., Kliewer, S.A., and Moore, D.D. (1997). Activation of the orphan receptor RIP14 by retinoids. Proc. Natl. Acad. Sci. USA *94*, 7909–7914.

GenBank Accession Number

The PXR sequence has been deposited in GenBank (accession number AF031814).

In re: Meyers
Appl. No. 09/464,039
Filed December 15, 1999

APPENDIX I

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 268, No. 29, Issue of October 15, pp. 22105-22111, 1993
Printed in U.S.A.

Identification of Regions in Interleukin-1α Important for Activity*

(Received for publication, April 28, 1993, and in revised form, June 23, 1993)

Richard B. Gayle, III‡§, Kurt Poindexter‡, David Cosman‡, Steven K. Dower‡, Steven Gillis‡, Thomas Hopp‡¶, Rita Jerzy‡, Shirley Kronheim‡, Vanessa Lum‖, Andrew Lewis‡, Marvin M. Goodgame‖, Carl J. March‡, Douglas L. Smith‖, and Subhashini Srinivasan‡

From the ‡Immunex Research and Development Corporation, Seattle, Washington 98101 and the ‖Life Sciences Research Laboratories, Eastman Kodak Company, Rochester, New York 14650

Saturation mutagenesis of the mature human interleukin-1α (IL-1α) gene has been performed. Following expression in *Escherichia coli*, the biological and receptor binding activities of the mutant proteins were examined. Most of the molecule could be altered with little effect on either function. More than 3,500 mutants were examined, and only 23 unique amino acid sequences were identified which resulted in an altered ratio of biological to binding activity when compared with wild-type IL-1α. These proteins possessed mutations at 38 of the 159 amino acid residues in IL-1α. Random mutagenesis at several of these positions identified further substitutions that affected activity. Examination of a model for IL-1α localized most of the residues which altered activity along one face of the molecule. This region appears to be distinct from areas of IL-1 which have been postulated to make contact with IL-1 receptor.

Interleukin-1 (IL-1)[1] is a potent cytokine that is involved in inflammatory responses and affects the growth and differentiation of T cells, B cells, and fibroblasts (for review, see Durum *et al.*, 1985). The two molecules responsible for this activity, IL-1α and IL-1β, share only 22% amino acid similarity (March *et al.*, 1985; Auron *et al.*, 1987). Each binds to both forms of IL-1 receptor (Dower *et al.*, 1986; McMahan *et al.*, 1991). Both IL-1 molecules are produced as intracellular precursors and are subsequently processed to mature proteins. Although the precursor form of IL-1α is biologically active, only the mature form of IL-1β has any biological activity (Mosley *et al.*, 1987a, 1987b). cDNAs have been isolated encoding a third form of IL-1, IL-1 receptor antagonist (IL-1ra) (Carter *et al.*, 1990; Eisenberg *et al.*, 1990; Hannum *et al.*, 1990). This molecule has homology to both IL-1α and IL-1β and has an affinity for the IL-1 receptors close to that seen for IL-1α and IL-1β, yet elicits no biological response from target cells (Arend *et al.*, 1990; Carter *et al.*, 1990; Eisenberg *et al.*, 1990; McMahan *et al.*, 1991). Alignments among these three sequences for several species do not indicate which residues are important for activity (Yanofsky and Zurawski, 1990). The three-dimensional structures of IL-1α (Graves *et al.*, 1990) and IL-1β (Gilliland *et al.*, 1987; Priestle *et al.*, 1990) demonstrate the structural similarity of the two molecules but do not suggest which regions of the molecules are responsible for activity.

Deletion and combinatorial mutagenesis have identified residues at the amino terminus of IL-1α which are needed for biological activity (Yanofsky and Zurawski, 1990). However, since the affinities of these mutants for IL-1 receptors were not examined, it is impossible to differentiate whether these mutations affect biological activity, the ability to bind IL-1 receptor, or the structural integrity of the protein. The existence of IL-1ra demonstrates the ability to separate biological activity from binding activity. Although many site-directed mutations of IL-1α and IL-1β have little effect on the function of the proteins (Gronenborn *et al.*, 1988; Kamogashira *et al.*, 1988a, 1988b; Craig *et al.*, 1989), several mutants demonstrate greatly reduced biological activity with little change in affinity for the type I IL-1 receptor (Gehrke *et al.*, 1990; Yamayoshi *et al.*, 1990).

Using a novel method of saturation mutagenesis, random mutations were generated throughout the entire sequence of IL-1α. Assays to determine both biological and binding activity were performed on several thousand mutant proteins. By examining the ratio of biological to binding activity for each mutant and comparing it with the ratio for wild-type IL-1α, regions of IL-1α which affect these two properties differentially were identified.

MATERIALS AND METHODS

Enzymes and Vectors—All restriction enzymes, T4 DNA ligase, and T4 polynucleotide kinase were obtained from Boehringer Mannheim or New England Biolabs. BBG1, a plasmid bearing a synthetic gene for human IL-1α, was purchased from British Biotechnology. The construction of pPLBBGIL-1α has been described previously (Poindexter *et al.*, 1991). An *Spe*I site was added using site-directed mutagenesis. Plasmid DNA was purified by the alkaline lysis method (Ausubel *et al.*, 1988).

DNA Synthesis—Oligonucleotide cassettes used for the construction of mutants were synthesized on an Applied Biosystems model 380A DNA synthesizer. For saturation mutagenesis each of the four phosphoramidites was contaminated with a small amount of the other three. The phosphoramidites were contaminated at two different levels, 4.2% for the sense strand and 8.4% for the antisense strand (Poindexter *et al.*, 1991). For random mutagenesis, the cassette was synthesized normally except for the substitution of an equimolar mixture of the four phosphoramidites for the three nucleotides making up the chosen codon. Oligonucleotides were purified by polyacrylamide gel electrophoresis on a 40-cm 8% polyacrylamide, 7 M urea gel. Care was taken to excise full-length oligonucleotides, and the

* This work was supported by funding from a joint venture between Immunex Corporation and Eastman Kodak. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) X55445.

§ To whom correspondence should be addressed: Immunex Research and Development Corp., 51 University St., Seattle, WA 98101. Tel.: 206-587-0430; Fax: 206-233-9733.

¶ Present address: Protein Research Laboratories, 10606-8 Camino Ruiz, Suite 281, San Diego, CA 92126.

[1] The abbreviations used are: IL-1, interleukin-1; IL-1ra, interleukin-1 receptor antagonist.

```
          ClaI (C)
                M  S  A  P  F  S  F  L  S  N  V  K  Y  N  F  M  R
1   atcgatactATGTCAGCACCTTTTAGCTTCCTGAGCAATGTGAAATACAACTTTATGAGG   60
              EcoRI (R)
    I  I  K  Y  E  F  I  L  N  D  A  L  N  Q  S  I  I  R  A  N
61  ATCATCAAATACGAATTCATTCTGAACGATGCATTGAACCAGTCTATTATTCGTGCAAAC  120
              PstI (P)
    D  Q  Y  L  T  A  A  A  L  H  N  L  D  E  A  V  K  F  D  M
121 GACCAGTACCTGACTGCAGCAGCCCTGCACAATCTGGACGAAGCAGTTAAATTCGACATG  180
           SstII (S)
    G  A  Y  K  S  S  K  D  D  A  K  I  T  V  I  L  R  I  S  K
181 GGTGCTTACAAGAGCTCGAAAGACGACGCAAAAATCACTGTAATCCTACGTATTTCTAAA  240
       PvuII (U)
    T  Q  L  Y  V  T  A  Q  D  E  D  Q  P  V  L  L  K  E  M  P
241 ACCCAGCTGTATGTAACTGCACAGGATGAAGATCAGCCAGTACTTCTGAAAGAAATGCCT  300
                 BamHI (B)
    E  I  P  K  T  I  T  G  S  E  T  N  L  L  F  F  W  E  T  H
301 GAGATCCCGAAGACTATCACTGGATCCGAGACTAACCTGCTGTTCTTCTGGGAAACTCAC  360
                 SpeI (S)
    G  T  K  N  Y  F  T  S  V  A  H  P  N  L  F  I  A  T  K  Q
361 GGTACCAAAAACTACTTCACTAGTGTGGCTCATCCGAACCTGTTCATCGCGACAAAACAG  420
              BglI (G)
    D  Y  W  V  C  L  A  G  G  P  P  S  I  T  D  F  Q  I  L  E
421 GACTACTGGGTATGCCTGGCAGGCGGTCCGCCATCGATCACTGACTTCCAGATCCTCGAG  480
                 HindIII (H)
    N  Q  A  *  *
481 AACCAAGCATAATAAgatctaagctt  506
```

FIG. 1. **Coding sequence of human IL-1α gene used in saturation mutagenesis.** The DNA sequence downstream from the promoter of pPLBBGIL-1α is shown (accession no. X55445). The coding sequence is represented by *capital letters*. The coding sequence for the IL-1α gene is divided into eight regions, named for the restriction enzymes that border them. The nine restriction enzymes are *Cla*I (C), *Eco*RI (R), *Pst*I (P), *Sst*I (S), *Pvu*II (U), *Bam*HI (B), *Spe*I (S), *Bgl*I (G), and *Hin*dIII (H). Thus the eight regions are CR, RP, PS, SU, UB, BS, SG, and GH.

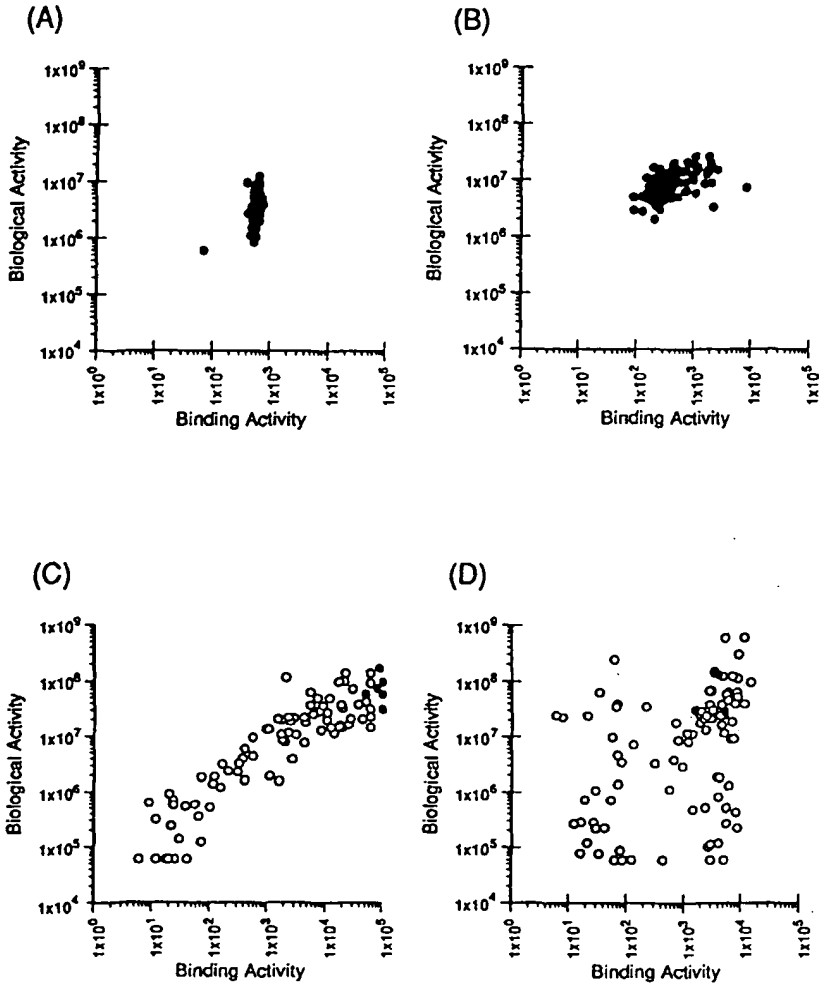


FIG. 2. **Ability of biological and binding activity screens to identify mutants with altered activity.** *Panel A*, 96 identical samples of wild-type IL-1α. *Panel B*, 96 independent samples of wild-type IL-1α. *Panel C*, 90 mutants of IL-1α from the PS region (*open circles*) and 6 wild-type IL-1α (*closed circles*). *Panel D*, 90 mutants of IL-1α from the RP region (*open circles*) and 6 wild-type IL-1α (*closed circles*). The biological activity is expressed in units/ml, and the binding activity is the reciprocal dilution that results in 50% binding inhibition (see "Materials and Methods").

oligonucleotides were deprotected and resuspended in TE (10 mM Tris-HCl, pH 7.4, 1 mM EDTA).

Assembly and Cloning of Mutagenic Oligonucleotides—The general procedure has been described (Poindexter *et al.*, 1991). Twenty picomoles of each oligonucleotide in a mutagenic cassette were mixed in 20 μl of TE and placed at 65 °C for 15 min. The mixture was allowed to cool slowly to room temperature and then placed on ice. Each mutagenic cassette had unique ends, allowing them to be ligated into appropriately cleaved vectors. To increase the efficiency of screening, intermediate vectors were constructed for each region. These intermediate vectors contained an irrelevant segment of DNA inserted between the relevant restriction enzyme sites. Insertion of the mutagenic cassette followed by restriction with an enzyme unique to the intermediate plasmid greatly reduced the incidence of vectors without

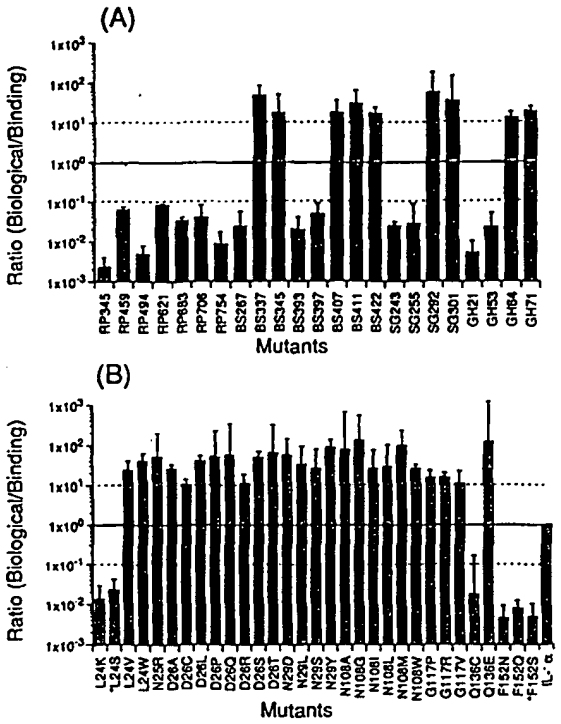


FIG. 3. **Mutants with altered activities.** The activity ratios were normalized to the internal wild-type IL-1α controls and the mean from at least three screens for mutants with normalized ratios 10-fold greater or less than IL-1α are shown. *Panel A*, proteins generated by saturation mutagenesis. *Panel B*, proteins generated by site-directed random mutagenesis at single amino acid residues. The single letter amino acid code is used to describe the amino acid in wild-type IL-1α, the number of the residue, and the amino acid of the mutant. An *asterisk* indicates single amino acid changes that were also isolated from the saturation mutagenesis screens.

a cassette inserted. The ligation mixture was transformed into GM1[pRK248]. The transformants were screened for insertion of the mutagenic cassette by colony hybridization (Poindexter *et al.*, 1991). Double-stranded sequencing of the vectors with inserts was performed using the dideoxy method (Sanger *et al.*, 1977).

Protein Expression and Analysis—Mutant proteins were produced by using a pH induction protocol (Poindexter and Gayle, 1991). Cells containing recombinant plasmids were inoculated into 24-well plates containing 1 ml of Superbroth (Ausubel *et al.*, 1988), supplemented with M9 minimal salts and 1% glucose. Following overnight growth at 30 °C, the pH of the medium was shifted to 9 by the addition of 5 M NaOH. The cells were grown at 30 °C at pH 9 for 18 h before the cells were pelleted. Cells (40 μl) were spun down in a 96-well plate at 3,000 rpm for 15 min. The cells were resuspended in an equal volume of lysis buffer (125 mM Tris, pH 8, 2% SDS) and then 80 μl of phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4 \cdot 7H_20$, 1.4 mM KH_2HPO_4, pH 7.4) was added. The samples were then used for biological or binding assays. Binding activity was determined by the capacity of *Escherichia coli* cell lysates to inhibit the binding of radiolabeled IL-1α to EL4 cells (Mosley *et al.*, 1987b). Biological activity was examined using an EL4 conversion assay (Mosley *et al.*, 1987a). *E. coli* lysates containing wild-type IL-1α were included in every assay as controls. The data were analyzed by a nonlinear least squares fitting routine, and activities were quantified using a standard curve, derived from purified IL-1α, for each biological assay and each binding assay.

Modeling IL-1α Structure—The three-dimensional model of IL-1α used in this study is an all atom protein model built using the Cα coordinates generated from the stereo diagrams of the crystal structure (Graves *et al.*, 1990). The *X* and *Y* coordinates for all of the Cα atoms were measured in the user unit using one of the two figures in the stereo diagram. The known standard distance of 3.8 Å, expected between the successive Cα atoms, was used to scale and assign the *Z* coordinates for all of the Cα atoms. An all atom model of 159 residues was constructed in FRODO (Jones, 1985) from these 151 Cα coordinates (Leu-7 to Asn-157). The structure was refined in a modified version of Biograf (Peeples and Goldstein, 1989).

Distances between the successive Cα atoms were computed in the *X-Y* projection by measuring the *X* and *Y* coordinates from one of the two stereo diagrams. A value of +1, −1, or 0 was assigned to the *Z* coordinates of every Cα atom if it was above, below, or in the same depth as the previous Cα atom in the stereo view. The scale factor required to change the user unit to Å unit was computed by equating the computed projection distance between two successive Cα atoms, with a sign of 0, to that of the standard distance of 3.8 Å. The *X* and *Y* coordinates of all of the Cα atoms were multiplied with this scale factor, and the projection distance between the successive Cα coordinates was computed. All of the projection distances measured this way will be either less than or equal to 3.8 Å because of the *Z* flattening. Appropriate *Z* coordinates were then assigned to each Cα atom such that the computed distance between successive Cα measured 3.8 Å. The computed *Z* coordinates were multiplied with the sign assigned to them. The cumulative error in the computation of the *Z* coordinates was corrected by adjusting only the *Z* coordinates using a wire model. The accuracy of the computed coordinate was tested by superimposing the resulting Cα trace on to the structurally conserved regions of IL-1β structure. The root mean square value for the structurally conserved region is 0.75 Å.

RESULTS

Alteration of the IL-1α Gene by Saturation Mutagenesis—The synthetic IL-1α gene has unique restriction enzyme sites approximately every 60 base pairs, dividing the gene into eight different regions (Fig. 1). The mutants from each region of the molecule are described by the restriction enzymes found at each end of the region. Thus the eight regions are CR (*Cla*I-*Eco*RI), RP (*Eco*RI-*Pst*I), PS (*Pst*I-*Sst*I), SU (*Sst*I-*Pvu*II), UB (*Pvu*II-*Bam*HI), BS (*Bam*HI-*Spe*I), SG (*Spe*I-*Bgl*I), and GH (*Bgl*I-*Hind*III). Saturation mutagenesis using cassettes for each region was performed throughout the entire molecule, relying on a technique that results in very low levels of wild-type sequences and roughly equal probabilities of one to five nucleotide changes in any one region (Poindexter *et al.*, 1991). The use of intermediate plasmids in the constructions, along with colony hybridizations (see "Materials and Methods"), greatly improved the yield of recombinant vectors, allowing insert frequencies over 90% to be achieved. More than 3,500 mutants were generated by this approach, encompassing the eight regions of the IL-1α gene (Fig. 1).

In one region (RP), 30 mutants were sequenced to determine accurately the mutation frequencies (Poindexter *et al.*, 1991). Those 30 mutants averaged 2.5 amino acid changes each, and every amino acid that could be altered in this region was found to be changed at least once. Several amino acids had three or four different substitutions. Sequencing mutants in each of the other seven regions did not reveal any deviation from the expected mutation frequency. More than 110 mutants were sequenced, and approximately 70% of the amino acid residues in IL-1α were altered at least once. At this rate of mutagenesis, screening approximately 520 mutants, or roughly 65 mutants from each of the eight regions, should result in a 99% chance that all of the amino acids in IL-1α which could be changed were altered at least once (for a discussion of calculating these probabilities, see Hutchison *et al.*, 1986). Thus, examination of 3,500 mutants should sample multiple mutations at every possible amino acid.

Screening of Mutant IL-1α Proteins—Determining both the biological activity and the ability to inhibit the binding of wild-type IL-1α to the type I IL-1 receptor for each mutant allowed molecules to be identified which affected these two functions differentially. The ratio of biological activity to binding ability was examined. This ratio represents an intrinsic specific activity of the molecule and therefore should be independent of protein concentration. To validate this ap-

RP Mutants	I	L	N	D	A	L	N	Q	S	I	I	R	A	N	D	Q	Y	L	T
RP345				D			M										D		
RP459					E		W												
RP494											M	P	S				N		
RP621					E					T									
RP683									L		S								
RP706					E						M								A
RP754/761			S																

BS Mutants	S	E	T	N	L	L	F	F	W	E	T	H	G	T	K	N	Y	F	T
BS267				T									D				F		
BS337					R							Q	V						
BS345				Y						K									P
BS393					P		S									K	F		
BS397								D				D		I					
BS407			A		I			Y									H		
BS411					K					D							D		
BS422							L			V		D							

SG Mutants	S	V	A	H	P	N	L	F	I	A	T	K	Q	D	Y	W	V	C	L
SG243				Y									L						
SG255									S								L		
SG292												T		H					
SG301				T								N						S	

GH Mutants	A	G	G	P	P	S	I	T	D	F	Q	I	L	E	N	Q	A
GH21										S							
GH53				L						V	H						
GH64						W	L					K					K
GH71					S		S										

FIG. 4. **Sequence of mutants that alter activity.** The amino acid sequence of the relevant cassettes is shown with the amino acid changes for each mutant underneath. The single letter amino acid code is used.

proach, the biological activity and binding activity of 96 identical samples of an *E. coli* lysate containing wild-type IL-1α were determined (Fig. 2*A*). Plotting the biological activity against the binding activity resulted in a cluster of points, with the greatest error being in the biological activity, presumably because of the greater inherent variation in this assay than in the binding assay. Screening *E. coli* lysates from 96 different inductions, each producing wild-type IL-1α, resulted in the greater scatter with a tendency for the points to cluster along a line whose slope equaled the activity ratio of wild-type IL-1α (Fig. 2*B*). Altering the concentration of wild-type IL-1α only moves the ratio along this line. Dilutions were performed to verify that the activity seen was linear with respect to IL-1α concentration (data not shown). The assays were able to distinguish levels of biological and binding activity over a range greater than 1,000-fold.

Mutants with a wild-type activity ratio should fall along the same line as wild-type IL-1α controls included in each assay. Proteins with increased biological activity in comparison to the amount of binding activity seen should fall above this line, whereas mutant proteins with decreased amounts of biological activity compared with the binding activity seen should fall below this line. For screening purposes, mutants were determined to have altered activity ratios if there was at least a 10-fold increase or decrease in the ratio relative to wild-type IL-1α for at least three independent sets of assays. In addition, mutants whose *E. coli* lysates lacked activity in either assay were rescreened to verify the lack of activity.

Approximately 1,700 clones were examined from the CR, PS, SU, and UB regions, encompassing more than one-half of the molecule. None of the mutants that demonstrated activity displayed any significant deviation from wild-type levels. These regions have an average of 2.2 amino acid changes per mutant (data not shown), which is not significantly different from the expected mutation rate. Fig. 2*C* displays a set of typical data for 90 mutants from the PS region of the molecule. The majority of these proteins have activity ratios similar to wild-type IL-1α. The other three regions give similar profiles.

Although most of the 1,800 mutants examined from the other four regions (RP, BS, SG, and GH) had activity ratios within 10-fold of wild-type IL-1α, each region included several mutants that had a ratio that deviated from the wild-type IL-1α ratio by more than 10-fold. An analysis of 90 typical mutants from the RP region is shown in Fig. 2*D*. There are several mutants in this group which have significantly different activity ratios from wild-type IL-1α. Although several of these mutants failed to maintain an altered activity ratio upon subsequent assays, many continued to have altered activity ratios.

Mutants from saturation mutagenesis with activity ratios 10-fold higher or 10-fold lower than wild-type are shown in Fig. 3*A*. IL-1α showed itself to be extremely resilient to change. Only 24 unique DNA sequences, out of more than 3,500 examined, produced protein that displayed a significant difference in activity from wild-type (Fig. 3*A*). This represents less than 0.7% of the mutants examined.

Of the 24 different mutants with altered ratios, 23 had unique amino acid sequences (Fig. 4). RP754 and RP761, although having the same amino acid sequence, have different DNA sequences. This demonstrates the power of this approach since it was possible to identify two independent clones

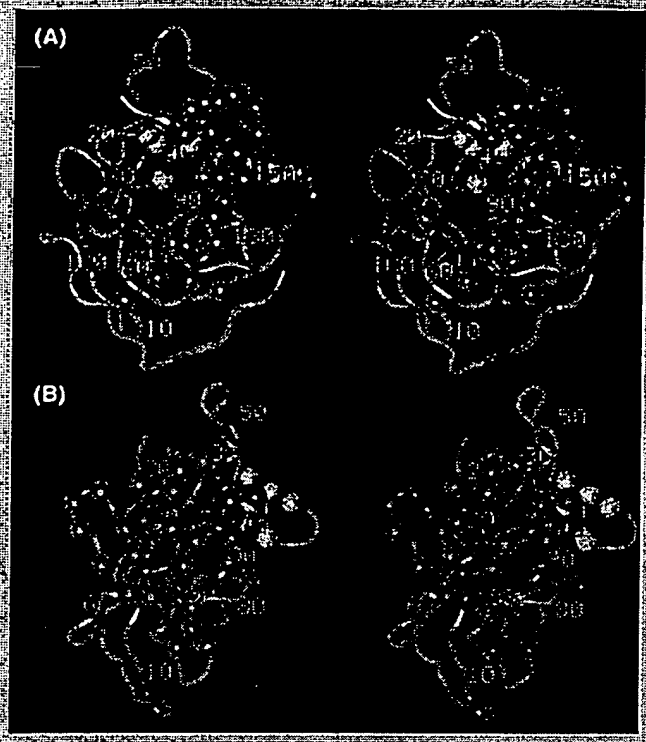


FIG. 5. Three-dimensional model of IL-1α. Stereoview showing position of residues that affect the activity of the molecule. *Panel A*, the plane of the figure is roughly perpendicular to the axis of the 12-stranded barrel. *Panel B*, rotated 90° from view in *panel A*. The α-carbon backbone is shown as a *ribbon*, with every 10th residue *numbered*. The *green spheres* illustrate the major region identified. The *red* and *yellow spheres* indicate the smaller regions made up of residues Glu-106, Asn-108, Leu-109, His-127, and Ile-33, Arg-34, Ala-35, Tyr-39, Thr-41, and respectively.

that had the same phenotype and the same amino acid changes out of a large pool of recombinants. The 23 distinct amino acid sequences contained mutations at a total of 38 amino acid residues. This corresponds to approximately 24% of the molecule.

Random Site-directed Mutagenesis—Because saturation mutagenesis often results in conservative amino acid changes, random mutagenesis was performed at several sites in the molecule. During the synthesis of a particular mutagenic cassette, a random sequence was placed at the codon of interest, allowing all possible amino acids to be produced at this position. Site-directed mutagenesis was performed at residues in the four regions identified by saturation mutagenesis. Leu-24 and Phe-152 were chosen because single amino acid changes introduced at these positions had already been found to alter the activity ratio (*i.e.* mutants RP754 and GH21). Asp-26, Leu-28, Asn-108, Gly-117, and Tyr-121 were examined because more than one mutant had been found which altered these residues. Asn-25 and Gln-136 were studied because they appeared to be conserved hydrophilic residues that were exposed to solvent, according to the crystallographic data (Graves *et al.*, 1990). Asn-29 was also mutated because it was one of the few conserved residues in the RP region of the molecule which was not identified using saturation mutagenesis. Up to 100 mutants at each position were examined. Screening proteins generated by random mutagenesis identified several further single amino acid changes that affected the activity ratios of the molecule (Fig. 3*B*). Although none of the changes at Leu-28 or Tyr-121 appeared to affect the ratio of biological activity to binding activity, substitutions at each of the other 8 amino acids produced at least one mutant that altered the activity ratio.

Spatial Location of Mutations—The crystal structures of IL-1α (Graves *et al.*, 1990) and IL-1β (Gilliland *et al.*, 1987; Priestle *et al.*, 1990) have been determined to 2.7 and 2.0 Å, respectively. Using a novel approach, the three-dimensional coordinates were determined from the published structure. The model of IL-1α superimposes on the similar trace from IL-1β with a root mean square difference of 0.75 Å. Residues that were determined by saturation mutagenesis and by random mutagenesis to affect activity ratios are displayed on the model for IL-1α (Fig. 5). Interestingly, these amino acids cluster in three regions. The majority are found along one side of the molecule, encompassing an area of approximately 600 Å^2.

DISCUSSION

More than 3,500 mutants were generated throughout IL-1α by saturation mutagenesis. The rate of mutagenesis was more than sufficient to produce several amino acid changes at every possible residue in these 3,500 mutants. The biological activity and the ability to inhibit the binding of IL-1α were measured for every mutant. The ratio of biological activity to binding activity gives a measure of the specific activity of each mutant. Antagonists will have low ratios, whereas mutants with high ratios demonstrate enhanced agonist activity.

Most of the molecule could be mutated with little effect on either activity. Combining the data from saturation mutagenesis and site-directed mutagenesis, alterations at only 39 positions resulted in proteins with modified activity ratios. Although mutations at these residues resulted in proteins with activity ratios up to 1,000-fold less than wild-type, it would appear that only a limited number of residues are critically required for activity. Most of the other 189 residues, or 75% of the molecule, may not contribute significantly to the biological activity of the molecule. As much as 68% of IL-1α may have little informational content, allowing a wide variety of amino acids to be substituted with little effect on activity (Zurawski, 1991). This is consistent with the observations reported in this paper.

Fig. 5 shows the spatial locations of the 39 residues changed in mutants with altered activity ratios. Almost all of the amino acid changes were found in β-strands, not in loops. With the exception of amino acids in strands 1 and 12, most of these amino acid residues have not been identified previously as important for activity. More than 75% of the identified amino acid residues are located along one face of the molecule. A substantial number of the mutated residues are located in β-strands 1, 2, 8, 9, 11, and 12. Several residues that may be involved in determining the activity of IL-1α, such as Asp-26, Lys-119, Gln-136, Ile-149, and Asp-151, appear in spacially similar positions in IL-1β.

In addition to this one major region there appear to be two smaller areas: one that includes Glu-106, Asn-108, Leu-109, and His-127, and another involving residues Ile-33, Arg-34, Ala-35, Tyr-39, and Thr-41. The former three amino acids form a small hydrophilic patch at the bottom and slightly behind the large region of mutated residues, whereas the latter amino acids form an exposed patch off to one side, separated from the main region by strands 3 and 4. The identification of three regions important for activity is intriguing. Since the IL-1 receptor is composed of three IgG-like domains, it has been postulated that each of the three domains interacts with a region on IL-1 (Clore *et al.*, 1991). Deletion of any of these three domains greatly reduces the binding of ligand (Dower and Sims, 1990).

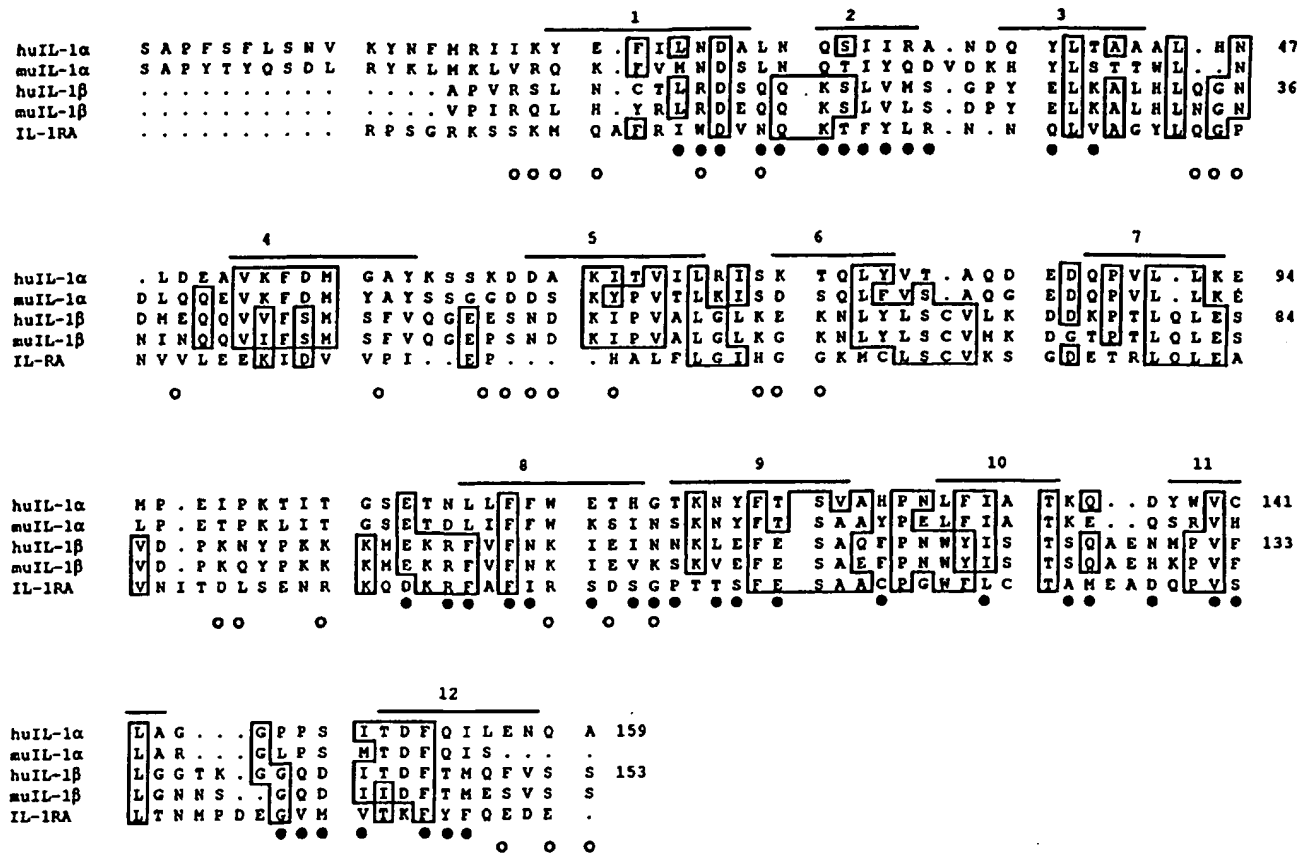


FIG. 6. **Alignment of human and murine IL-1α, human and murine IL-1β, and IL-1 receptor antagonist.** Residues that are conserved in at least three of the sequences are *boxed*. The β-strands for IL-1α are shown above the sequence. The numbering for human IL-1α and IL-1β is displayed on the *right*. *Closed circles* indicate amino acid residues that were changed in IL-1α mutants with altered phenotypes. *Open circles* indicate residues in human IL-1β which have been postulated to interact with the type I IL-1 receptor.

Using site-directed mutagenesis, single amino acid changes that decrease or increase the ratio of biological activity to binding ability were found at several positions in the molecule. At most of these locations only the substitution of specific amino acids had any impact on activity. Screening up to 100 isolates revealed only a small number of specific amino acid substitutions at Leu-24, Asn-25, Asn-29, Gly-117, Gln-136, and Phe-152 which altered the activity ratio. Substitutions at Asp-26 and Asn-108 were exceptions. The presence of any of 10 amino acids at Asp-26 altered the activity ratio. Changes at Asp-26 which reduce the amount of biological activity have been reported (Yamayoshi *et al.*, 1990). This may indicate that an Asp residue at position 26 is required for wild-type activity. Several hydrophilic substitutions at Asn-108 did not appear to affect biological activity (Zurawski, 1991). However, placing hydrophobic residues at this position did affect the ratio, with a relative increase in the biological activity.

Few other reports of IL-1α mutagenesis have distinguished between changes that affect biological activity and those that also alter the interactions of the protein with the receptor. Mutations at Asp-151 differentially affect biological activity and binding ability (Yamayoshi *et al.*, 1990). Examination of a large group of IL-1α mutants has previously revealed the importance of Asp-26 and Asp-151, but only the effects on biological activity were studied (Kawashima *et al.*, 1992). Random combinatorial mutagenesis has identified amino acids in which substitutions have no effect on biological activity (Yanofsky and Zurawski, 1990). The lack of mutations at certain residues was used to infer the importance of these residues for biological activity. There was, however, no demonstration of the ability of these mutants to interact with the IL-1 receptor. Since only biological activity was examined, it is difficult to determine whether the constraints on amino acid residues were functional in nature or whether there were structural constraints as well.

Sequence alignments of human and mouse IL-1α and IL-1β are shown in Fig. 6. These molecules bind to the type I IL-1 receptor and are active in the EL4 conversion assay. Inclusion of IL-1ra, a molecule that also binds to IL-1 receptor but produces no biological activity, in the alignment does not immediately reveal any obvious region that is responsible for the uncoupling of biological activity and the ability to bind receptor. The 39 amino acids identified in this report are also shown. Although many of these amino acids are found toward the COOH terminus of the molecule (which is the most highly conserved region of the molecule), there does not appear to be much selection for conserved residues. In fact, regions in which few amino acids are conserved, such as in strand 2 or strand 8, have several residues that appear to be important for activity. Although there are no data demonstrating that IL-1α and IL-1β interact with the type I IL-1 receptor in the same fashion, residues of IL-1β which have been shown to be important for biological activity are found in positions homologous to some of the residues identified for IL-1α (Gehrke *et al.*, 1990; Ju *et al.*, 1991).

Interestingly, residues of IL-1β which have been shown to

influence binding to the type I IL-1 receptor or have been proposed to interact with the receptor (Clore *et al.*, 1991; Grenfell *et al.*, 1991; Labriola-Tompkins *et al.*, 1991; Veerapandian *et al.*, 1992) are generally located outside the regions of IL-1α shown in Fig. 5. Out of 45 amino acids postulated to be involved in binding of IL-1β only 9 overlap with residues identified in this study (Fig. 6). IL-1α and IL-1β either have different regions interacting with the receptor, or there is a large region in IL-1α and IL-1β which is required for biological activity but not for high affinity interactions with the receptor.

A superfamily of molecules with protein folding similar to IL-1 has been proposed (Graves *et al.*, 1990; Murzin *et al.*, 1992). This superfamily includes certain proteinase inhibitors and heparin-binding growth factors, such as fibroblast growth factor. Two forms of fibroblast growth factor have been shown to fold in a very similar fashion to IL-1α and IL-1β (Ago *et al.*, 1991; Eriksson *et al.*, 1991; Zhang *et al.*, 1991; Zhu *et al.*, 1991), displaying 12 β-strands with a pseudo 3-fold symmetry. Several regions of the fibroblast growth factor molecule have been identified which are important for activity. These areas occupy spatially similar regions of the fibroblast growth factor three-dimensional structure as the amino acids of IL-1α identified by saturation mutagenesis. The modes of interaction between members of the IL-1 superfamily and their respective receptors may involve similar regions of the folded protein.

The manner in which IL-1 generates a biological response is complicated. There are two different ligands, an IL-1 antagonist and two forms of IL-1 receptor (Dower *et al.*, 1990). There is evidence for multiple pathways of signal transduction (for review see Sims *et al.*, 1993). The effect of IL-1 on a particular cell type may depend on the receptor found on the cell and which signaling pathway is being used. Analysis of the effect of these mutations on the different biological responses may be helpful in further elucidation of IL-1 signal transduction.

REFERENCES

Ago, H., Kitagawa, Y., Fujishima, A., Matsuura, Y., and Katsube, Y. (1991) *J. Biochem.* **110**, 360–363

Arend, W. P., Welgus, H. G., Thompson, R. C., and Eisenberg, S. P. (1990) *J. Clin. Invest.* **85**, 1694–1697

Auron, P. E., Warner, S. J., Webb, A. C., Cannon, J. G., Bernheim, H. A., McAdam, K. J., Rosenwasser, L. J., LoPreste, G., Mucci, S. F., and Dinarello, C. A. (1987) *J. Immunol.* **138**, 1447–1456

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. J., Smith, J. A., and Struhl, K. (1988) *Current Protocols in Molecular Biology*, John Wiley and Sons, New York

Carter, D. B., Deibel, M. J., Dunn, C. J., Tomich, C. S., Laborde, A. L., Slightom, J. L., Berger, A. E., Bienkowski, M. J., Sun, F. F., McEwan, R. N., Harris, P. K. W., Yem, A. Y., Waszak, G. A., Chosay, J. G., Sieu, L. C., Hardee, M. M., Zurcher-Neely, H. A., Reardon, I. M., Heinrikson, R. L., Truesdell, S. E., Shelly, J. A., Eessalu, T. E., Taylor, B. M., and Tracey, D. E. (1990) *Nature* **344**, 633–638

Clore, G. M., Wingfield, P. T., and Gronenborn, A. M. (1991) *Biochemistry* **30**, 2315–2323

Craig, S., Pain, R. H., Schmeissner, U., Virden, R., and Wingfield, P. T. (1989) *Int. J. Pept. Protein Res.* **33**, 256–262

Dower, S. K., and Sims, J. E. (1990) in *Cellular and Molecular Mechanisms of Inflammation* (Cochrane, C. G., and Gimbrone, J. M. A., eds) pp. 137–172, Academic Press, New York

Dower, S. K., Kronheim, S. R., Hopp, T. P., Cantrell, M., Deeley, M., Gillis, S., Henney, C. S., and Urdal, D. L. (1986) *Nature* **324**, 266–268

Dower, S. K., Bomzstyk, K., and Sims, J. E. (1990) *Prog. Clin. Biol. Res.* **349**, 241–249

Durum, S. K., Schmidt, J. A., and Oppenheim, J. J. (1985) *Annu. Rev. Immunol.* **3**, 263–287

Eisenberg, S. P., Evans, R. J., Arend, W. P., Verderber, E., Brewer, M. T., Hannum, C. H., and Thompson, R. C. (1990) *Nature* **343**, 341–346

Eriksson, A. E., Cousens, L. S., Weaver, L. H., and Matthews, B. W. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 3441–3445

Gehrke, L., Jobling, S. A., Paik, L. S., McDonald, B., Rosenwasser, L. J., and Auron, P. E. (1990) *J. Biol. Chem.* **265**, 5922–5925

Gilliland, G. L., Winborne, E. L., Masui, Y., and Hirai, Y. (1987) *J. Biol. Chem.* **262**, 12323–12324

Graves, B. J., Hatada, M. H., Hendrickson, W. A., Miller, J. K., Madison, V. S., and Satow, Y. (1990) *Biochemistry* **29**, 2679–2684

Grenfell, S., Smithers, N., Witham, S., Shaw, A., Graber, P., and Solari, R. (1991) *Biochem. J.* **280**, 111–116

Gronenborn, A. M., Wingfield, P. T., McDonald, H. R., Schmeisaner, U., and Clore, G. M. (1988) *FEBS Lett.* **231**, 135–138

Hannum, C. H., Wilcox, C. J., Arend, W. P., Joslin, F. G., Dripps, D. J., Heimdal, P. L., Armes, L. G., Sommer, A., Eisenberg, S. P., and Thompson, R. C. (1990) *Nature* **343**, 336–340

Hutchison, C. A. I., Nordeen, S. K., Vogt, K., and Edgell, M. H. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 710–714

Jones, T. A. (1985) *Methods Enzymol.* **115**, 157–171

Ju, G., Labriola-Tompkins, E., Campen, C. A., Benjamin, W. R., Karas, J., Plocinski, J., Biondi, D., Kaffka, K. L., Kilian, P. L., Eisenberg, S. P., and Evans, R. J. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 2658–2662

Kamogashira, T., Masui, Y., Ohmoto, Y., Hirato, T., Nagamura, K., Mizuno, K., Hong, Y. M., Kikumoto, Y., Nakai, S., and Hirai, Y. (1988a) *Biochem. Biophys. Res. Commun.* **150**, 1106–1114

Kamogashira, T., Sakaguchi, M., Ohmoto, Y., Mizuno, K., Shimizu, R., Nagamura, K., Nakai, S., Masui, Y., and Hirai, Y. (1988b) *J. Biochem. (Tokyo)* **104**, 837–840

Kawashima, H., Yamagishi, J.-i., Yamayoshi, M., Ohue, M., Fukui, T., Kotani, H., and Yamada, M. (1992) *Protein Eng.* **5**, 171–176

Labriola-Tompkins, E., Chandran, C., Kaffka, K. L., Biondi, D., Graves, B. J., Hatada, M., Madison, V. S., Karas, J., Kilian, P. L., and Ju, G. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 11182–11186

March, C. J., Mosley, B., Larsen, A., Cerretti, D. P., Braedt, G., Price, V., Gillis, S., Henney, C. S., Kronheim, S. R., Grabstein, K., Conlon, P. J., Hopp, T. P., and Cosman, D. (1985) *Nature* **315**, 641–647

McMahan, C. J., Slack, J. L., Mosley, B., Cosman, D., Lupton, S. D., Brunton, L. L., Grubin, C. E., Wignall, J. M., Jenkins, N. A., Brannan, C. I., Copeland, N. G., Huebner, K., Croce, C. M., Cannizzarro, L. A., Benjamin, D., Dower, S. K., Spriggs, M. K., and Sims, J. E. (1991) *EMBO J.* **10**, 2821–2832

Mosley, B., Dower, S. K., Gillis, S., and Cosman, D. (1987a) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 4572–4576

Mosley, B., Urdal, D. L., Prickett, K. S., Larsen, A., Cosman, D., Conlon, P. J., Gillis, S., and Dower, S. K. (1987b) *J. Biol. Chem.* **262**, 2941–2944

Murzin, A. G., Lesk, A. M., and Chothia, C. (1992) *J. Mol. Biol.* **223**, 531–543

Peeples, W. J., and Goldstein, P. (1989) *Cytobios* **58**, 109–123

Poindexter, K., and Gayle, R. B., III. (1991) *Gene (Amst.)* **9**, 125–130

Poindexter, K., Jerzy, R., and Gayle, R. B., III. (1991) *Nucleic Acids Res.* **19**, 1899–1904

Priestle, J. P., Schar, H. P., and Grutter, M. G. (1990) *Prog. Clin. Biol. Res.* **349**, 297–307

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467

Sims, J. E., Bird, T. A., Giri, J. G., and Dower, S. K. (1993) *IL-1 Signal Transduction: Signal Transduction through Growth Factor Receptors* (Kitagawa; Y., ed), Jai Press, Greenwich, in press

Veerapandian, B., Gilliland, G. L., Raag, R., Svensson, A. L., Masui, Y., Hirai, Y., and Poulos, T. L. (1992) *Proteins Struct. Funct. Genet.* **12**, 10–23

Yamayoshi, M., Ohue, M., Kawashima, H., Kotani, H., Ilda, M., Kawata, S., and Yamada, M. (1990) *Lymphokine Res.* **9**, 405–413

Yanofsky, S. D., and Zurawski, G. (1990) *J. Biol. Chem.* **265**, 13000–13006

Zhang, J. D., Cousens, L. S., Barr, P. J., and Sprang, S. R. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 3446–3450

Zhu, X., Komiya, H., Chirino, A., Faham, S., Fox, G. M., Arakawa, T., Hsu, B. T., and Rees, D. C. (1991) *Science* **251**, 90–93

Zurawski, G. (1991) *Trends Biotechnol.* **9**, 250–257

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.